



**GEOGRAPHIC INFORMATION SYSTEM  
AS A MAP AND SURVEY DATABASE  
FOR A SELECTED AREA**

by

**JACOBUS HENDRIK RAUBENHEIMER**

Thesis presented in partial fulfilment of the requirements for

the degree of

Master of Science in Engineering (M.Sc.(Eng.))

Department of Surveying and Geodetic Engineering

University of Cape Town

November 1995

The University of Cape Town has been given  
the right to reproduce this thesis in whole  
or in part. Copyright is held by the author.

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

## **ACKNOWLEDGEMENTS**

I would like to express my sincere appreciation to the following people:-

My supervisor Mr Colin Martin for his guidance and support, Mr Gabri Rigotti of Compu Stowe Design for valuable advice with the database and program design, and Mr Pine Pienaar for proof reading the document.

The staff of the Directorate of Surveys and Land Information, the Surveyor-General in Cape Town and to André van den Berg of Wellington Municipality, all who have assisted me by providing data.

My wife Ronélle and children Pieter and Anneke who have missed a husband and father during the months I've worked on this research, but still supported me.

The glory go to our God who gives us the knowledge and the ability to do our work so that it can be to His glorification.

## **DECLARATION BY CANDIDATE**

I hereby declare that this thesis is my own work and that it has not been submitted for a degree at any other University.

Signed by candidate

J H Raubenheimer

Cape Town

15 November 1995

## SYNOPSIS

The purpose of this research was to identify how the distribution and availability of spatial data could be improved. This should then minimize the duplication of data and ensure a better utilization of available data sources.

All decisions that are made should be based on information, and especially decisions about our natural resources should be based on geographical information - this is spatial information of our environment. Many users, however, are not aware of the spatial data available or where to find it.

In Chapter 2 the literature review reports on the spatial data sources in other countries, and different methods that are used to make this data available to the users. Most of the research was done on the distribution of digital data, and not much on traditional non-digital material. The establishment of national land information systems and data sharing via the networking of databases is receiving wide attention.

An investigation was done to establish the spatial data suppliers and the data users, the types of data supplied, and the requirements. This was done by means of a questionnaire survey and personal interviews. Data suppliers and users were grouped into four categories:

- *Spatial data managers* - they are the major suppliers of spatial data and include air survey companies, cartographers and surveyors. Spatial data supply is their main function and they produce mostly topographic, cadastral, political/administrative and road data.
- *Planners and construction* - this group consist of engineers, town planners and architects and are the major users of spatial data. They normally use the basic data and enhance it by adding information. They have established links with the major suppliers and know where to obtain the required data.
- *Land use and development* - these include ecologists, geologists, farmers and property developers and are mostly users of spatial data. They use a large variety of data, e.g. topographical, cadastral, vegetation, soil, geologic and meteorological. They are also producers of some unique data such as the distribution of certain bird species or the potential and soil type of agricultural land.
- *Services and information* - this is provided by libraries, educational institutions, Telkom and the CSIR.

It was found that there is a good supply of data and the shortcomings are rather due to uninformed users - they are not always aware of the variety of data available.

Three possible methods were investigated to improve the distribution of spatial data and to reduce the duplication of data:

- A comprehensive GIS with a full database.
- A national spatial data infrastructure (NSDI).
- A metadata base.

A comprehensive GIS is ideal when an organisation has the hardware and expertise to operate and make use of the analysis functionality. In an environment where many users do not have access to the hardware, this would not provide access to spatial data for them.

A national spatial data infrastructure which creates a network between different databases, has the same disadvantages and would only operate effectively amongst users who have similar hardware and database installations. Only 40% of the organisations and individuals in the "land use and development" and "services and information" groups were able to make use of digital data. A GIS or NSDI would therefore not be accessible to the other 60%.

A metadata base stores data about data, it provides background information on the nature of data, the sources from which it was derived and the overall quality of the material. A metadata base provides "pointers" to other databases and also to other forms of spatial information. A metadata base was found to be the most suitable spatial information system to provide for the requirements of all the data suppliers and users because of its versatility.

The principles of database design were used to develop a relational database to store the metadata. This was done to ensure data consistency and to avoid redundancy. This relational

database would be accessed via a standard GIS.

A geographical database was developed to serve as a base map for the alphanumeric data. The cadastral land parcel was taken as the basic spatial entity which would be used for queries to the database. Geographical data such as roads, railway lines and rivers were shown for orientation purposes only.

Different methods of data retrieval were investigated and it was found that a menu system was the most suitable. It is proposed that the agency operating the metadata base should be paid a commission when a sale is made by the spatial data supplier to the user. This would then encourage all parties to share data instead of duplicating it.



**TABLE OF CONTENTS**

<b>ACKNOWLEDGEMENTS</b> .....	<b>ii</b>
<b>SYNOPSIS</b> .....	<b>iv</b>
<b>TABLE OF CONTENTS</b> .....	<b>viii</b>
<b>LIST OF ILLUSTRATIONS</b> .....	<b>xiii</b>
<b>GLOSSARY</b> .....	<b>xiv</b>
<b>1. INTRODUCTION</b> .....	<b>1</b>
1.1 INFORMATION .....	1
1.2 AVAILABILITY OF SPATIAL DATA .....	2
1.3 DATA SHARING .....	3
1.4 GEOGRAPHIC INFORMATION SYSTEM .....	3
1.5 PURPOSE OF THE STUDY .....	5
1.6 RESEARCH METHOD .....	6

2. LITERATURE REVIEW .....	7
2.1 INFORMATION .....	7
2.2 SPATIAL DATA .....	12
2.2.1 Spatial Data Sources .....	14
2.2.2 Digital Data .....	16
2.3 DEMAND FOR SPATIAL DATA .....	20
2.4 GEOGRAPHICAL DATABASE .....	24
2.4.1 Database Design .....	24
2.4.2 National Land Information System (NLIS) .....	29
2.4.3 Data Sharing .....	33
2.4.4 Database Index .....	37
2.4.5 Metadata .....	42
2.4.6 Temporal Database .....	44
2.5 MAP LIBRARY .....	46
3. INVESTIGATION ON AVAILABLE SPATIAL DATA .....	52
3.1 METHOD OF INVESTIGATION .....	52
3.2 SPATIAL DATA SOURCES .....	56
3.2.1 State .....	58
3.2.2 Statutory Bodies and Utilities .....	63
3.2.3 Private Organisations .....	65
3.3 USERS OF SPATIAL DATA .....	67
3.3.1 Spatial Data Managers .....	68
3.3.2 Planners and Construction .....	69

3.3.3 Land Use and Development . . . . .	70
3.3.4 Services and Information . . . . .	71
3.3.5 Politicians and Entrepreneurs . . . . .	71
3.3.6 General Public . . . . .	72
3.4 CONCLUSIONS ON THE AVAILABILITY OF SPATIAL DATA . . . .	73
4. DESIGN OF GIS DATABASE . . . . .	76
4.1 INTRODUCTION . . . . .	76
4.1.1 Data Management . . . . .	76
4.1.2 Philosophy of the Database Design: Affordability and Simplicity . . . . .	79
4.2 FORMAT OF DATA . . . . .	80
4.2.1 Extensive Geographic Information System . . . . .	80
4.2.2 National Spatial Data Infrastructure . . . . .	82
4.2.3 Metadata Base . . . . .	84
4.2.4 Conclusions . . . . .	85
4.3 DATABASE PLANNING . . . . .	88
4.3.1 Design Process . . . . .	88
4.3.2 Analysis of Information Needs . . . . .	89
4.3.3 Information Model . . . . .	91
4.4 REQUIREMENTS ANALYSIS . . . . .	92
4.4.1 Scope of the Database . . . . .	92
4.4.2 User Views . . . . .	93
4.5 CONCEPTUAL DESIGN . . . . .	94

4.5.1 Data Modelling . . . . .	94
4.5.2 View Integration . . . . .	98
4.5.3 Conceptual Data Model . . . . .	99
4.6 IMPLEMENTATION DESIGN . . . . .	104
4.6.1 Mapping to a Logical Model . . . . .	104
4.6.2 Designing Applications . . . . .	104
4.6.3 Physical Design . . . . .	105
5. BASE MAP AND INTEGRATION OF ATTRIBUTE DATA . . . . .	107
5.1 REQUIREMENTS . . . . .	107
5.2 BASE MAP . . . . .	109
5.2.1 Contents . . . . .	109
5.2.2 Acquisition . . . . .	113
5.3 INFORMATION INTEGRATION . . . . .	114
5.3.1 Definition . . . . .	114
5.3.2 Integration of Spatial and Attribute Data . . . . .	115
5.4 UPDATING THE GEOGRAPHICAL DATA . . . . .	117
5.4.1 Cadastral Data . . . . .	117
5.4.2 Topographic Data . . . . .	118
6. RETRIEVAL OF METADATA . . . . .	119
6.1 INTRODUCTION . . . . .	119
6.2 REQUIREMENTS . . . . .	120
6.2.1 Diversity of Data . . . . .	121

6.3 DATA RETRIEVAL .....	123
6.3.1 Dialogue Styles .....	123
6.3.2 Output .....	125
6.4 DESIGN CRITERIA .....	127
6.4.1 Menu Design .....	128
6.5 ACQUIRING METADATA .....	129
7. CONCLUSIONS .....	131
7.1 SPATIAL DATA SUPPLY AND DEMAND .....	131
7.2 DATA DISTRIBUTION .....	133
7.3 METADATA BASE .....	134
7.4 DATA SHARING .....	135
BIBLIOGRAPHY .....	137
APPENDIX A: QUESTIONNAIRE ON SPATIAL DATA .....	157
APPENDIX B: ORGANISATIONS OR INDIVIDUALS THAT COMPLETED THE QUESTIONNAIRE ON SPATIAL DATA .....	167

## LIST OF ILLUSTRATIONS

Figure 2.1	Decision-making structure	9
Figure 2.2	A hub network model for a NLIS	30
Figure 2.3	Nodal approach to a land database configuration	41
Figure 3.1	Categories of spatial data suppliers and users	55
Figure 3.2	Major types of spatial data supplied. Percentages between brackets are when supply for own use are included.	57
Figure 3.3	Demand for different types of spatial data	67
Figure 3.4	Number of respondents that supply or use digital data	68
Figure 3.5	Types of spatial data supplied to the users	72
Figure 4.1	The components of a NSDI	83
Figure 4.2	Typical steps in database design	89
Figure 4.3	Relationships between entities	91
Figure 4.4	A supplier supplies many types of spatial data to many users	92
Figure 4.5	Metadata required by map user	93
Figure 4.6	Metadata required for an orthophoto map	93
Figure 4.7	The unnormalized relation 'Map'	95
Figure 4.8	The relation 'Map' in 1NF	96
Figure 4.9	The relation 'Map' in 2NF	97
Figure 4.10	The relations 'Map_Data' and 'Map_Type' are in 3NF	98
Figure 4.11	A data model of the relations for maps and for aerial photographs	100
Figure 4.12	A many-to-many relationship is implemented by creating a new relation 'Erf_Map'	101
Figure 4.13	Database tables and the relationship between them	103

## GLOSSARY

**Attribute** - a descriptive characteristic of a feature, e.g. what it is, age, cost, owner, and how to display it: colour, length, symbol, etc.

**Automated cartography** - the process of drawing maps with the aid of computer driven display devices such as plotters and graphics screens. The term does not imply any information processing.

**Base map** - a map portraying background reference information onto which other information is placed. Base maps usually show the location and extent of natural features and permanent man-made objects.

**Cadastral system** - is a method of securing ownership of land and other real rights in land. It consist of two components: spatial and numerical data which represent the actual beacons demarcating the land boundaries and are registered in the Surveyor-General's office, and a legal document, the deed which is registered in the Deeds office. Together these two documents define a property and its owner.

**Catalogue** - a list of maps or other items, arranged in some definite order. It records, describes and indexes the resources of a collection, a library or a group of libraries.

**Database** - a collection of interrelated information, usually stored on some form of mass-storage system such as magnetic tape or disk, and accessed using the facilities of a database management system (DBMS). A GIS database includes data about the position and the attributes of geographical features that have been coded as points, lines, areas, pixels or grid cells.

**Database management system (DBMS)** - a software system with facilities for database language processing, to permit both the handling of run-time calls for database access from application programs and/or end-users, and the maintenance of data integrity.

**Digital elevation model (DEM)** - a DEM is a three-dimensional model depicting a part of the earth surface in a digital format. This data is usually represented in a regular grid in which the planimetric position of each point is implicit, and the elevations are represented in raster form. (Clarke 1995)

**Digital map** - the representation of cartographic features in a form that allows the values of their attributes to be stored, manipulated, and output by a computer system. A digital map is a database or file that becomes a map when a hard copy or screen display is produced.

**Entity** - a person, place, object or concept about which an organisation chooses to store data.

**Entity-Relationship (E-R) data model** - a type of network data model that uses a special symbol (a diamond) to represent associations between entities.



**GIS** - a geographic information system for capturing, storing, checking, integrating, manipulating, analysing and displaying data which are spatially referenced to the earth. This is normally considered to involve a spatially referenced computer database and applications software.

**Information** - structured data which can be used for decision making. The following operations may be performed on information: create, transmit, store, retrieve, receive, copy, process, destroy.

**Land information system (LIS)** - refer to systems containing land ownership information, usually consisting of the numerical (spatial) and legal data of the cadastre. This is generally regarded as a subset of a GIS.

**Layers** - a logical separation of mapped information according to theme. Many geographic information systems and CAD/CAM systems allow the user to choose and work on a single layer or any combination of layers at a time.

**Map** - it is a medium for the comprehension, recording and communication of spatial relationships and forms. It contains information such as location, direction, distance, height, connectivity, contiguity, adjacency, hierarchy and spatial association.

**Menu** - a simple method for easy command selection whereby the user make a choice, either by entering a number or letter or via the cursor control keys.

**Metadata** - the documents and records used to provide background information on the nature, contents and format of the data, the sources from which it was derived and the overall quality of the material. This is valuable for those intending to use the data for deriving further information. Other information may include procedures of data capture, accuracy, types of analysis performed and rules for the cartographic display of the data. (Cassettari 1993: 43)

**National Land Information System (NLIS)** - is a computerised management system established to provide decision-makers access to spatial information. It would maintain a list of suppliers and users and an index of data. (Lester 1993: 2)

**National Spatial Data Infrastructure (NSDI)** - a national network linking databases and users and enhancing the accessibility, communication and use of geographically referenced data. Its purpose is to enable the electronic transport of data, as proposed by McLaughlin and Nichols (1994: 62).

**Normalization** - the process of decomposing complex data structures into simple relations according to a set of dependency rules. (McFadden & Hoffer 1988: 670)

**ReGIS** - a vector based commercial geographic information system developed in South Africa by Automated Methods (Pty) Ltd. It can access files in a number of commercial database management systems, e.g. dBASEIII, Sybase, Oracle, Informix and Adabas.

**Relational database** - the data is structured as a table, in which each row of values (tuple)

corresponds to a logical record, and the column headings are the names of the fields (items) in the records. Using the convention of normalization, each tuple contains data representing the properties either of a "real-world" entity or the relationship between two or more entities. In itself, a set of relations leaves the relationship between relations implicit; these can be made explicit in, for instance, an entity-relationship diagram.

**Spatial data** - data or information with implicit or explicit information about location.

**Spatial meta-information system (SMIS)** - a system that makes it possible to retrieve meta-information on the basis of a spatial reference. (Medyckyj-Scott 1991: 86)

**SQL** - Structured Query Language. A standard data definition and manipulation language for relational databases.

**Topology** - the relationships between different spatial objects (points, lines, polygons), e.g. which polygons share a boundary. Finding a path through a network is a topological calculation.

## **CHAPTER 1**

### **INTRODUCTION**

#### **1.1 INFORMATION**

It has been said that we are living in the age of the information explosion. The basic resource in all decision making is information. In practice, many decisions are made on the basis of inadequate information, in a disjointed and incremental way, and for reasons that are often subjective. The availability of good information can prevent neither mismanagement nor the taking of wrong decisions. It can, however, reduce the level of ignorance of the consequences of action or inaction. Information is different from other commodities: it does not get used up, it can lose value even if it is never used, the same information can be used by many users, and it is inexpensive to reproduce. Information is vital, especially for the management of our natural resources and the environment. Dale and McLaughlin (1988) have noted:

"Without accurate information about the lands and waters, and without an up-to-date inventory of the country's resources and what is happening to them and to the environment, the government and the people are handicapped in controlling their own destiny. It is not possible to make best use of the land and natural wealth, or to prevent its misuse, without good factual knowledge of the country and its features."

## 1.2 AVAILABILITY OF SPATIAL DATA

Spatial data and geographical information are related terms, and both can be related to specific locations on the earth. It covers an enormous range, including the distribution of natural resources, descriptions of infrastructure, patterns of land use and the health, wealth, employment, housing and voting habits of people.

It is often difficult to discover whether spatial data on a particular topic already exist and, if so, where it is held, who owns it and how it may be accessed. Much of the difficulty stems from the complexity of the data structures typical of spatial data in general, or GIS in particular. They are not easily summarized in a simple verbal statement, and practical GIS specialists working on projects have little time or motivation to fill in detailed questionnaires about the data. This difficulty is now compounded by the volume of the data that have already been collected.

Walker *et al.* (1992) showed that the recognition of the need for data users to locate and retrieve useful data has led to the establishment of a number of what are

variously called data directories, inventories or catalogue systems.

### **1.3 DATA SHARING**

To promote the use of spatial data and to minimize the duplication of data, the sharing of data must be encouraged. This can be done in various ways, but the most important is to inform the users of what data is available, where is it available and in what form.

A culture of data-sharing must also be developed amongst data suppliers who often tend to be possessive of their data. They should realise that sharing data is also to their benefit because it would increase the market for their products.

### **1.4 GEOGRAPHIC INFORMATION SYSTEM**

It is difficult to define a geographic information system (GIS) because some people believe that hardware and software are the central focus, others believe that information processing or even applications is the key element. One fact is however certain: it is a multi-disciplinary science, involving cartographers, social scientists, planners, geographers, and many others. One definition of a GIS which is widely accepted, is that it is a computerised database management system for the capturing, storing, retrieval, manipulation, analysis and display of spatial data. GIS generally consist of four generic components: hardware, software, spatially related data, and

management procedures; all of which are required to use a GIS for management, analysis or decision making in a spatial domain. The alphanumeric and locational data of features are equally important, and it is the combination of these that makes a GIS unique. Topology, or intelligence connected to features, sets it apart from computer aided design (CAD) systems, and the difference with a standard database is that in the GIS, features have geographical positions.

According to Goodchild (1990), there appear to be four views underlying the applications of GIS: (1) GIS as automated mapping; (2) GIS as map analysis; (3) GIS as inventory; (4) GIS for spatial analysis and spatial decision support. It is clear that the full benefit of GIS is only used by the last.

The main reason for the rapid development of GIS technology is the rapid advances in computer hardware. At present there are five major computer hardware approaches to supporting a GIS: mainframe computers, minicomputers, workstations, microcomputers and computer networks. Microcomputers are the fastest growing hardware platform for GIS because of their low cost, increasingly powerful computational capabilities and ability to be networked together. GIS software comprises basically two subsystems: the first manages cartographic data, in either raster or vector formats; the second, a relational database management system (DBMS), is used to manage the attribute data. They are tightly related to one another so that an alteration in one is immediately followed by a corresponding alteration in the other.

The goals of using GIS technology were identified by Hiland *et al.* (1991) as the improvement of communication between managers and scientists; to simplify and promote the integration of data from different sources; to reduce the duplication of data and effort; to identify the most important data sets available and to promote data exchange.

## 1.5 PURPOSE OF THE STUDY

This study will investigate different methods of improving the distribution of spatial data. It will investigate who are the suppliers of spatial data and what type of data they supply, who are the users of spatial data and what are their requirements.

It will show that spatial data consist of many different types, e.g. digital and hard copy maps, photographs, orthophoto maps, digital imagery, co-ordinates and textual data. The fact that much of this is not in a digital format must be taken into account.

The main objectives of the study are:

- to investigate the suppliers and users of spatial data and their requirements;  
and
- to develop a method of spatial data distribution so that it can be used with the minimum of data duplication. This should be to the benefit of both data



suppliers and users.

## **1.6 RESEARCH METHOD**

A questionnaire survey and personal interviews were used to determine who are the data suppliers and users and the type of data they supply or use. Different methods were then evaluated that could be used to distribute the spatial data. A metadata base was developed containing data about spatial data to enable the users to determine which data would fulfil their requirements.

## **CHAPTER 2**

# **LITERATURE REVIEW**

### **2.1 INFORMATION**

Everybody is constantly making decisions about important or minor issues. Whether it is a major political decision on the foreign policy of a country, a business decision on the buying of shares, or a personal decision about which restaurant to visit for dinner, such a decision may have favourable or disastrous results.

The decision-making process is controlled by a number of factors:

- the regulations which define the constraints within which the decision is made;
- the various and changing circumstances pertaining at the time;
- an individual's experience of similar decision-making situations; and

- knowledge about the desired outcomes.

(Cassettari 1993)

Decision-making is to weigh the various alternatives carefully, address the relative importance of the various factors and apply considerable judgement which best meet the objectives. To do this we need suitable and sufficient information. Four kinds of information can be identified.

- *Data* is unstructured and presents the information in its barest and most elementary form, usually as a string of characters. By itself, an item of data is not of much use and must be combined with other data to become meaningful and useful.
- *Text* or information is structured and processed in a specific format so that it can be used to make meaningful decisions.
- *Images* present information in pictorial form, e.g. charts, graphs, maps and freehand drawings.
- *Voices* refer to spoken phrases and would also include recordings.

Within a company decisions are made on a day to day basis by managers and workers. It is necessary to delegate decision-making to the level at which the appropriate information is available and where the resulting action can be taken.

Cassettari (1993) illustrated this process as shown in Figure 2.1.

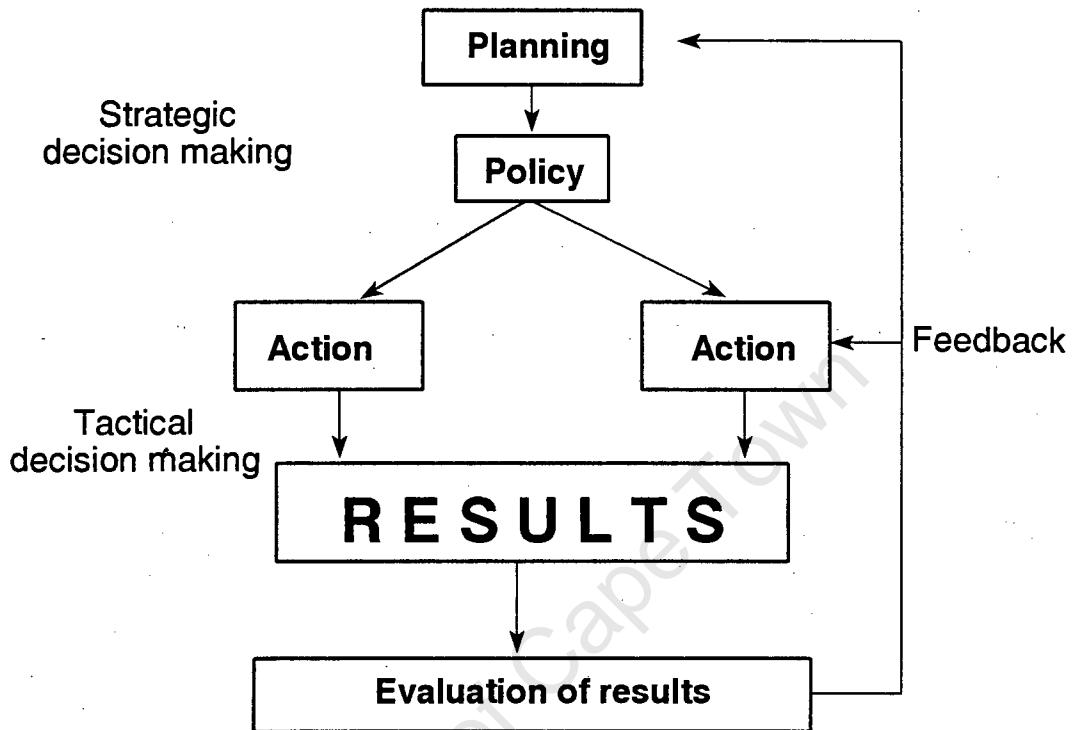


Fig. 2.1 *Decision-making structure*

To aid the decision maker in deriving acceptable solutions, computer-based tools have been developed that can evaluate large volumes of data and utilise modelling functions to help identify relationships and trends within the data. One example is decision support systems (DSS) which are used to support operational research in finding solutions to business problems. Densham (1991) set out six characteristics of a decision support system:

- They are explicitly designed to solve ill-structured problems where the objectives of the decision-maker and the problem itself cannot be fully or precisely defined;

- they have a user interface that is both powerful and easy to use;
- such systems enable the user to combine analytical models and data in a flexible manner;
- they help the user to explore the solution in space (the options available) by using the models in the system to generate a series of feasible alternatives;
- they support a variety of decision-making styles and are easily adapted to provide new capabilities as the needs of the user evolve; and
- such systems allow problem solving to be both interactive and recursive - a process in which decision-making proceeds by multiple paths, perhaps involving different routes, rather than a single linear path.

If spatial or geographical data is to be used in the decision-making process, the concept can be expanded to make provision for a spatial decision support system (SDSS). Because of the nature of complex spatial problems, however, a SDSS will need to provide additional capabilities and functions that

- provide mechanisms for the input of spatial data;
- allow representation of the complex spatial relations and structures that are common in spatial data;

- include analytical techniques that are unique to both spatial and geographical analysis; and
- provide output in a variety of spatial forms including maps and other, more specialized, types.

Land is the foundation of all forms of human activity. From it we obtain the food we eat, the shelter we need, the space to work in, and the room to relax. In both the private and public sectors, land information is a prime requisite for making decisions related to land investment, development, and management. Information reduces uncertainty by helping to identify and analyse problems. Strategies to overcome them, may then be prepared and implemented. This land information is provided by geographers, surveyors, cartographers, foresters, valuers, and others who have traditionally played a leading role in the land information field, as well as systems engineers, computer scientists, records managers, town planners, lawyers and resource specialists (Dale and McLaughlin 1988). Clarke (1989) pointed out that land related information is scattered amongst a number of organisations, there is duplication of information implying also duplication of effort, which makes it difficult to find the most suitable information at the best price at the right time.

A geographic information system can assist us in gathering the relevant information about land, evaluating it in a consistent and objective manner to come to meaningful decisions and actions. A good information system is one that provides us with the necessary data relevantly organised so that we can make the right decision about the

real world.

## 2.2 SPATIAL DATA

Spatial data may take many forms (Laurini and Thompson, 1992):

- *real*, for example the terrain conditions or buildings;
- *captured*, that is, recorded by physical devices like electronic sensors and cameras;
- *interpreted*, that is, involving some human intervention as in field sketches of landscapes or a questionnaire;
- *encoded*, as in paper maps or digital data; and
- *structured* or organised in some way, such as tables in census reports or data in geographic information systems.

Traditionally hand-drawn maps were used to display the arrangement of features on the earth's surface. In the last century two inventions changed the technology of map making: the camera and the aeroplane. But the map products continued to look much the same and contained essentially the same categories of information. Computer

technology and especially the ability to store spatial data have resulted in major changes in the production of maps. The map, as a primary tool for meeting spatial or geographical information needs, is being supplemented and displaced by new computer-based tools that have come to be known as geographic information systems (GIS).

In a society which depends increasingly upon access to information, effective and timely means of dissemination are at a premium. Mapping organisations are also under financial pressures. They may be forced to recover a specified portion of their costs from revenues on products and services. Most mapping organisations are undergoing a shift in emphasis from printed maps, charts and atlases to the supply of digital data.

Thapa and Burtch (1991) argued that data collection is by far the most expensive part of establishing a GIS and more emphasis should be paid to accuracy, age and quality of the data. They recommended classifying the collection of spatial data as primary or secondary. Primary methods refer to data collection either directly from the field or from recent aerial photographs and satellite imagery. This can be classified under geodesy (control), surveying, photogrammetry and remote sensing. Secondary methods refer to data collection from existing sources such as maps, charts, documents, etc. These may be classified as manual (digitizing), semi automatic (line scanning) and automatic (scanning).



### 2.2.1 Spatial Data Sources

Fisher (1991) distinguished between analogue and digital data sources. Maps, as a major source, have a number of limitations: (1) a paper map has a specific scale, whereas a digital map is scale independent, and due to their size certain objects may be excluded from the base map; (2) the type of map user determines the type of information included and how it is presented; (3) currency - there is a time delay between data gathering and map publication; (4) map coverage - source maps for an area are not always complete or available at the same scale; (5) map accuracy may vary, this include positional accuracy and accuracy (correctness) of the attribute data; and (6) map sheets and series - edge-matching between sheets may be a problem. The major digital data sources are remote sensing, GPS (with dataloggers for attribute data) and ground surveying with electronic theodolites and dataloggers.

Remote sensing, and particularly photogrammetry, is a major data source for GIS. Since 1972 with the launch of LANDSAT-1, high resolution digital remotely-sensed data have become widely available. Because the imagery is in digital form, the data can be computer-processed to directly generate geographical information. Surface temperature, land use, crop condition, flooding, water quality and forest harvesting are a few examples of geographical data that can be produced using computer methods. Computer enhancement of digital imagery enables other resource information to be

visually interpreted and used to update the information in a GIS.

Every country has a national mapping organisation responsible for producing the official maps for general use. The United States Geological Survey (USGS) provides digital cartographic data, maps, aerial photographs, satellite image data and geodetic control information to its clients (Thorley *et al.* 1993). To keep pace with the budget they intended to combine resources with other spatial data producers. This would include the exchange of funds (or joint funding), information and products. This would also reduce data duplication. To ensure acceptable standards, all suppliers have to adhere to the spatial data transfer standard. This should maximise the usefulness and accessibility of the data to multiple users:

According to Rhind (1993) the clients in Britain require the following information from Ordnance Survey (OS):

- a 'seamless' database of information derived from the OS large scale mapping;
- an up-to-date record of the postal addresses and a grid reference for each and every property in Britain;
- provision of the basic map series in hard copy and digital format; and

- transformation of the National Grid to be compatible with position-fixing obtainable from the Global Positioning System (GPS).

Adlam *et al.* (1988) found that the need for geoscience data is largely for detailed 'raw' data or interpretations of aggregated data (such as maps), preferably in a form which can be used immediately.

Sörbom and Svensson (1990) described the juridical, economic and technical information contained in the Swedish land data bank system. This is divided into three major groups:

- the real property register with information about designation, area, co-ordinates, legal surveys, plans and regulations;
- taxation information with information about land use, assessed value, etc.; and
- the land register with information concerning ownership, lease-holder, mortgages, etc.

### 2.2.2 Digital Data

In a report on spatial data needs in the United States of America Bossler *et al.* (1990) made some far-reaching recommendations, amongst others that

a "data donor" program to a national digital cartographic database be created. A national exchange standard should be provided to all state bodies to ensure that spatially-referenced digital products are compatible with the national database, also with respect to accuracy and quality. Incentives, especially to data donations from the private sector would include data exchange agreements and work-shared and cost-shared programs. They also recommended that a digital data quality standard be established, and that a national spatial database be created. This should be feature-oriented and accessible on line by the year 2010, if not sooner. This database should also refer users to specialized data sets created outside the national database, but registered to it. It should also include provisions for systematic update of features and data layers.

If a convenient and timely supply of needed items is not available, the client reluctantly utilises an alternative product, no matter how inferior or unsatisfactory it may be. The result of this is the loss of a sale and, even more importantly, an unhappy customer whose perception of the service offered, and of the publisher, is sadly diminished. McArthur and McGrath (1991) stated that there is a need for a comprehensive brochure on available spatial data provided by national mapping organisations to provide an inventory of products and services and to provide a mechanism whereby services and products can be requested. They stated that it was unlikely that information on digital data could be provided adequately in such a brochure because the target markets are too diverse, knowledge of digital data is not

yet comparable to that of analogue products, copyright may exist on the data and detailed information would be required on the area of interest and storage medium for digital data.

Fisher (1991) identified the data problem but also indicated a possible direction for the solution: "Discovering geographical data sources is a problem in many countries, but in some a centralized GIS makes an obvious point of contact." This was also shown by Eyre (1989) and Cassettari (1994) who proposed a "spatial data hypermarket" where GIS users could browse, compare, select and purchase analogue and digital data of all sorts.

Laursen (1990) described a method whereby six organisations in the Faroe Islands jointly established a land information system. They represent a national survey organisation, an engineering office, a telephone and electricity company and two local authorities. The collection of information from different organisations raised the following questions:

- who is responsible for misuse of correct data?
- who is responsible for correct use of erroneous or incomplete information?
- what about unauthorized use of data?

- what about financing and distribution of costs?

He stated further that "in this model it is a must to work together on a common solution, because otherwise the society cannot afford the expenses necessary to establish a system like this". Difficulties encountered were incompatibility of the different databases and how to ensure that all map updates get to users. A standard exchange format is therefore essential.

Greve *et al.* (1993) pointed out the problems encountered with the recording of historical information. In the past this was done by retaining copies of previous versions of hard copy maps. In future this will require a capability to capture the history of individual features within the database. Langran (1992) pointed out that a temporal GIS would trace the changing state of a study area, storing historic and anticipated geographical conditions. It could respond to the following queries:

- where and when did change occur?
- what types of change occurred?
- what is the rate of change?
- what is the periodicity of change?

From this, the software might assess:

- whether temporal patterns exist;

- what trends are apparent; and
- what processes underlie the change.

## 2.3 DEMAND FOR SPATIAL DATA

Spatial data requirements may vary: to plan a route we would need co-ordinate information of the roads, addresses, names of locations, important features that can be used as landmarks and street names; to analyse the impact of chemical nutrients that are discharged into streams, we would need surface gradients and aspects, the sequence of land uses along the river courses, soil permeability, underground hydrology and type of vegetation.

We normally have two categories of spatial problem solving:

- choosing a location to meet some requirements, e.g. the best school site; or
- inferring the form of objects not directly observable, e.g. to determine the geological composition of the soil.

Dymon (1989) conducted research to determine the profile of map users. Although the research investigated the users of groundwater maps in the USA, the conclusions should be similar for map users in general. He found that private sector employees had a technical orientation that enabled them to respond well to the technical information on the maps, while the public sector employees lack such an orientation and thus may experience problems when using the maps. The investigation also

showed that people in the public sector are mainly involved with policy-oriented work, those in the private sector are mainly involved in engineering (specific to the map application) and had formal training in a related discipline. The group from the private sector knew the types of data they wanted from the maps and where to obtain the maps. Morrison (1979) investigated the frequency of map use by drivers, especially for route-planning. As could be expected the map use was in direct proportion to the annual mileage driven. What was, however, significant is that map use also increased with the level of education and when the age group is nearer to 40.

The demand for spatial data and the analysis of this data have largely assisted the development of GIS, and all the main fields in the spatial environment have also been involved in its development: cartography, computer science, geography, surveying, remote sensing (including photogrammetry), commercial data processing, mathematics and statistics (Coppock and Rhind, 1991). These were used for environmental protection, urban and regional planning, land management, property ownership and taxation, resource management, transportation planning, the management of utilities, site location, military intelligence and tactics, etc.

According to Anderson and Callahan (1990) it is the intention of the U.S. Geological Survey to have a digital database by the year 2000, therefore some development tasks are being implemented to: (1) expand and improve mass digitization capabilities; (2) modify data structures to support increased content and access requirements; (3) develop digital revision capability; (4) develop product generation capability for standard, derivative, and digital products; (5) improve quality control; and (6) support



advanced analysis and applications. The spatial data of the USGS database are in general obtained by digitizing quadrangle maps, which was considered to be the most practical approach. This database would also be scale-specific, therefore suitable for use within a given graphic scale range. It should include the following categories of data: public land survey system, boundaries, geodetic control, transportation systems, hydrography, cultural features, vegetation, land use and ownership. Eventually updated information will be obtained from other sources such as aerial photography, other imagery or field data.

To populate a database, data will invariably be obtained from many different sources. Flowerdew (1991) described the problems involved in data integration, i.e. the process by which different sets of data within a GIS are made compatible with each other. These may include the quantitative measure of the data or the size of whatever is being studied; the representation of an observation, whether by a point, line or an area and how that is located; the time the data was collected; and the measurement and locational errors. He concluded: "The ability to combine together data of many different types and to display them in any combination is the main factor differentiating a GIS from mere database management systems on one hand and computer mapping systems on the other, it is the very heart of GIS".

Aronoff (1989) noted the important elements of data quality: positional accuracy, attribute accuracy, relational accuracy between data elements, resolution, completeness, time (of mapping), lineage of a data set, i.e. its history, the source data and processing steps used to produce it. To ensure positional accuracy Donahue

(1990b) pointed out the importance of basing survey data on a national geodetic control system. Although this is done for cadastral surveys in South Africa it also applies to all other survey data. He added: "If the information we create today is not quality information, it will not be demanded in the future. If it is not collected today, then it will have to be obtained and maintained in the future."

According to Maffini (1990) and Rhind (1992) there are five factors influencing the distribution of government databases: political, legal, fiscal and administrative, economic and moral. A traditional view is that the public has already paid for the collection of data by the state through taxes, the information already used for these purposes should, it is argued, be made generally available at no more than the cost of reproduction. It is also held that low cost dissemination maximizes the breadth of use and thus facilitates the creation of taxable wealth and jobs. Another argument is that only a small number of citizens may benefit from the free availability of data which has been paid for by all and that this is unfair. It is also regarded that a price on information inevitably leads to more efficient operations and forces consumers to specify exactly what they require. In the United States maps and digital data are charged at a level to cover the full cost of reproduction and distribution - but not to make a profit. There are also no restrictions on the copyright of these data. Most other countries add some value to cover part of their production costs (Rhind 1992, 1993). Because the value of data degrades with time, the use and therefore also the price decreases with time; however this does not apply to historical data.

For most information retrieval tasks, one does not simply ask the system a question

and have it respond with all relevant information and nothing but relevant information. To use an information system (IS), a query language such as Structured Query Language (SQL) is normally used. This could be done via keyboard, mouse trackerball or touchscreen. It is argued that the best way is through pointing (a menu system) rather than through typing. Medyckyj-Scott (1991) sets out the requirements for a query on an IS or a spatial meta-information system (SMIS):

- an interface that is orientated to the user;
- a means of querying the system that is easy to learn, usable and flexible;
- hard-copy output;
- the ability to store queries for future use;
- a task-orientated help facility; and
- a facility to enable depositors of data and managers to enter and edit data easily and on-line.

## **2.4 GEOGRAPHICAL DATABASE**

### **2.4.1 Database Design**

A database should be independent of its applications and its design should concentrate upon the data that is actually collected and stored. If this philosophy is accepted, the danger of producing the wrong type of data for future applications can be greatly reduced. Some special requirements of

users from a digital database are:

- it must be possible to obtain only a selection of the data set;
- it must be possible for users to add their own data;
- regular updates and date given should be provided to ensure currency of data; and
- personalised data retrieval tools (menus) should be provided for users who do not have the expertise to extract the data that will suit their requirements.

The differences between the objectives of producers and users of digital data can be distinguished. Producers want to continue to supply data with the same type of information and general form as traditional maps. Users claim that this does not make full use of digital technology, thus not taking full advantage of digital data. It also does not allow them to exploit computer analysis to the full if the data structures are not suitable.

Emphasis should be placed on the following:

- the type of data collected should be what the user requires and in a suitable format;
- the amount of data should be sufficient for analysis to be performed with it; and
- the data structure should support all applications in an efficient

manner.

Coleman and McLaughlin (1994) defined spatial data infrastructure as "encompassing the data sources, systems, network linkages, standards and institutional issues involved in delivering spatially-related information from many different sources to the widest possible group of potential users". They suggested that an information infrastructure should possess the following characteristics: standardized contents (data), conduit (telecommunications network) and flow-control procedures; tight (preferably electronic) connections between major suppliers and users of information; and database and communications components customized for easy third-party access.

As for any information system, the database management system is essential for the proper functioning of a GIS. Frank (1988) set the following requirements:

- selection of data for retrieval should be based on a multitude of access keys (e.g. name of person, address);
- provide standardized access to data and separation of data storage and retrieval functions from programs using the data (this makes the database and application programs independent, so that changes in the one do not necessarily lead to changes in the other);

- interface between database and application programs based on a logical description of the data (details of the physical storage structure should be transparent to the applications);
- make access functions in applications independent of the physical storage structure, so adaptations to expanding storage needs do not influence the application programs;
- allow for access to the data by several users at the same time;
- provide for the definition of consistency constraints for the data which will then be automatically enforced, where consistency constraints are rules which must hold for all data stored, and are an excellent technique to reduce the number of errors in a large data collection; and
- access to the data should be possible both from a high level language and from a user-friendly query language.

Special attention must be paid to fast access of spatial data which are quite different from alphanumeric data. Data protection is extremely important.

According to McLaughlin and Nichols (1994) the management considerations for a national spatial data infrastructure (NSDI) are:

- developing appropriate organisational structures;
- providing effective leadership in and among organisations;
- developing strategies for managing private sector involvement;
- allocating rights and responsibilities, including the custodianship of data and the responsibilities for the maintenance of the databases and networks;
- strategic planning and project management, including the design of implementation strategies;
- promoting and marketing the NSDI and, eventually, its services;
- accounting and other financial arrangements; and
- developing directories and other common services.

They pointed out that during the 1980s spatial information systems were mainly used for textual data and in simple, straightforward query applications of the spatial data. "No organisation wanted to be left out, but few have had innovative ideas about using the system they purchase." As we approach the 21st century the emphasis is shifting from information technology to information and its use. Even information will become part of the background as communicating and extracting knowledge become increasingly important.

#### **2.4.2 National Land Information System (NLIS)**

The objective of a NLIS is to provide a meaningful, co-ordinated and integrated computer-based land information system on a national basis.

By meaningful is meant that the NLIS will provide, as far as possible, all the information required by the users. By co-ordinated is meant that the establishment and maintenance of the NLIS is done in an orderly manner with co-operation between all the concerned organisations. By integrated is meant that it must be possible to combine two or more data sets with the minimum of effort and therefore it should conform to certain predetermined standards.

Sabel and Ralphs (1992) described the proposed National Land Information System for Britain (known as the Domesday 2000 project) and its relationship with the Land and Property Gazetteer (Pearman 1993) and the Street Gazetteer. Instead of including all the data sets in one enormous database, the preferred model is a 'hub' network illustrated in Figure 2.2. This 'hub' network would operate by having central controlling software, which individuals could access via online terminals. The data would be held locally by the organisations owning the information. By using networking technology to link the individual entities, it is hoped to provide a seamless user interface, which would have the advantages of controlling security and pricing centrally, whilst enabling each information supplier to maintain



ownership of the data at their remote site. Using this model, one could ensure that the data was kept up-to-date, which is one of the overriding principles of the project, since it is accurate, up-to-date information which is of most value to the user.

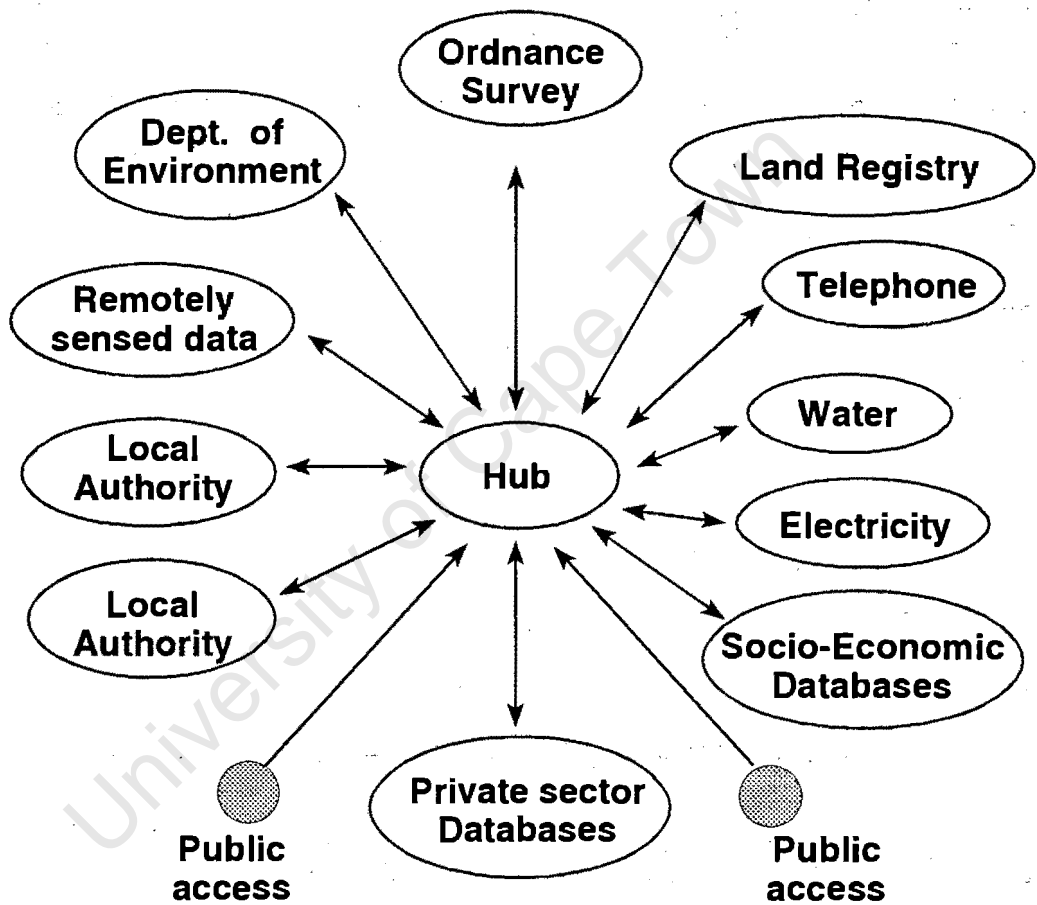


Fig 2.2 *A hub network model for a national land information system (after Sabel & Ralphs 1992).*

The first priority, however, was to determine whether there is a viable market for the product and to determine the kinds of information that users would require. According to Smith (1994) the primary objectives of the NLIS pilot project are:

- to perform a data analysis of property data maintained by the Valuation Office, the Land Registry and the Ordnance Survey and to create a pilot data set for a local authority;
- to convert this pilot data set to a single conformant database (i.e. a Land and Property Gazetteer); and
- to develop and implement a pilot system to gain information on the operation of a NLIS based on the Land and Property Gazetteer.

It is interesting to note that a large component of this research is to determine a definition for a basic land and property unit and unique property reference numbers. South Africa's cadastral and registration system will make this a relatively easy problem to solve.

According to Grant (1993) the purpose of the Australian digital cadastral database is to:

- create administrative efficiency and eliminate the duplication of equipment, staff, expertise and time;
- improve the availability of land related data;
- facilitate public access to information;
- enable the development of new, high quality land information products;
- facilitate effective management of the land resource by providing the foundation for spatially referenced data in an interactive database

system;

- assist all levels of government, public sector agencies and the business community to make better management decisions; and
- generate revenue for government.

Provision was also made for the incorporation of natural resources that would:

- reduce cost to government by eliminating duplication and the sharing of resources between agencies involved in the collection of natural resource information;
- improve quality of decisions related to natural resources;
- provide a fundamental information resource for policy formulation; and
- improve coordination between public sector agencies resulting in more efficient operations.

This is similar to the South African National Land Information System which is to provide all the information (as far as possible) required by the users (Clarke 1989 and Lester 1993). The major application would be in planning in all its different forms and to provide spatial information to all possible users. For a data set to be included it should be considered a natural resource required by two or more users. The Coordinating Committee for the NLIS identified the following entities as a minimum to

be on the NLIS: roads, rivers, cadastral boundaries, deeds (ownership), population distribution, urban land use, utilities, bridges, dams, demography, buildings and railways.

The major supplier organisations are the Directorates of Control Surveying and Mapping and the offices of the Surveyors-General, the offices of the Registrars of Deeds, Central Statistical Services and local authorities, and, with the exception of the local authorities, have all expressed their willingness to take on this responsibility.

#### **2.4.3 Data Sharing**

Often large organisations build up data resources piece by piece, an approach which can lead to inefficiency and duplication. Kuggeleijn (1995) listed the example of the Netherlands Ministry of Transport, Public Works and Water Management all of which have on occasion, bought identical data sets from the same supplier several times over. This emphasizes the importance of spatial data management. Strater (1992) listed three policy issues regarding the distribution of digital data:

- Who can have a copy of the data? Issues such as public records, copyright and confidentiality of data must be considered.
- What can the end-user do with the data? A supplier of data may provide a disclaimer which documents information on data quality,

scale and source materials.

- Should the end-user pay for the data? Providing data free of charge promotes data sharing. Other options are paying for data transfer only or providing services in kind.

Sussman (1993) showed that large corporations had a huge mass of data stored many times by many individuals. In the city of Scarborough a total of 125 data sets existed which, in total, required 28 staff years per annum to maintain. Most of these data sets were actually components of data stored elsewhere in the organisation with the addition of only a small portion of new data. It was clear that greater sharing of data was essential.

The city of Newport Beach, California, provides access to its GIS to its citizens via a terminal at the library (Corbley 1995). These users can be divided into three broad groups. The first are new residents who want an overview of the city, e.g. the name and address of their councillor. The second group are home builders and developers seeking land parcel information required for permits. A permit application can be completed and submitted at the computer terminal, with some of the information already filled in. The primary users are the local business community using the demographic information to include in their marketing and business plans. Population and commercial growth statistics are provided for each neighbourhood.

An automated system for searching and retrieving geographical, cartographic and bibliographic data by linking original programming with an existing GIS software package and a standard relational database management system was described by Hiland *et al.* (1991). It provide the capability for users to search for data references by defining the area on a displayed map. The system will then return any of the following: a list of digital maps or imagery that can be displayed immediately and visually overlaid, a list of maps/remotely sensed data and information on their availability, and a list of bibliographic references concerning the area and subject defined. Selections are available for performing bibliographic and geographical searches for textual material as well as maps, satellite imagery, photography, video and survey data via a menu system. Three options are provided for access to the data. The GIS either stores a reference to the data and the name of a contact person, or access to the owner's computer is obtained via a network, or the data sets are physically stored on the GIS facility.

Price (1995) defined the ideal situation for data sharing: Ideally a computer network would link all users and producers of spatial data. As data was edited, the new data would instantly be available to everyone else on the network. Some problems that will have to be overcome are:

- a large volume of data is not yet available in digital form;
- data producers are often unwilling to supply data, due to a reluctance

to take responsibility for the data and because of concerns of pricing and copyright;

- it is often difficult for users of digital data to find out what data is available and where to get it;
- the telecommunication links for wide-area networks are not always available;
- data users with a thorough understanding of networking and database technology are scarce; and
- in most cases data producers use incompatible data models, feature coding and formats.

The conditions for a successful data exchange system are:

- an organisation must be put in place to co-ordinate the development of standards by users and producers of data, and to maintain the data catalogue, and data definition dictionary;
- each database must have a standard structure, and all the contents must conform to this standard;
- all data to be shared must have a common set of data definitions agreed by all producers and users;
- a data catalogue containing full information on all available data must be kept up-to-date and circulated to all data users; and
- data should be available in a system-independent format.

Such a national LIS may take several forms, e.g.

- A network with databases at data producers' nodes which are accessible from any point on the network. This is only practical when the database structures at each site are essentially the same.
- A central database holding all common data and regularly updated by producers.
- A central database holding a catalogue of the available data and its location. This provides central control for accounting purposes, data access, and for the enforcement of standards.

Price concluded that the main problem is not technical but human organisation and management. A major advantage would be to eliminate duplication of data capture.

#### **2.4.4 Database Index**

Cuthbertson (1993) found that a database which incorporates the access to numerous databases would be cumbersome and difficult to manage and maintain, especially if the interfaces were diverse. Therefore a "pointer" to them would be required. He recommended the creation of a database with two parts, one with the metadata and the other with the data to be analysed.

A number of researchers actually took this further and showed that it was



not viable to store all the data sets as part of the main database, but rather to provide an index or bibliography to the data sets. This is especially the case where a vast amount of data in a large number of data sets have to be made available. A good example of this is the Louisiana Coastal GIS Network described by McBride *et al.* (1991) of which the purpose is to implement a system for information access that provides references and sources to spatial data and that links existing GIS databases, and to organise these diverse, multi-agency data sets into an automated access system. Furthermore, this data comes in a wide variety of media types, including digital maps, high-resolution seismic profiles, aerial videotape surveys, satellite imagery, tabular records, aerial photography, and field surveys. Unfortunately, much of the digital data is unknown to other agencies or inaccessible due to storage medium, data structure, or hardware platform.

Objectives that are similar to that of the author of this paper are to:

- improve communication between scientists using spatial data, such as planners, engineers, surveyors, cartographers, etc.;
- simplify the integration of spatial data from a variety of sources;
- eliminate duplication of effort so funds can be spent more efficiently;
- identify the most important databases in existence and incorporate them into the GIS network as digital information;
- develop guidelines for cataloguing different media (maps, photographs, satellite imagery, videotape surveys, textual attributes)

and establishing data set ancestry; and

- develop a user interface that provides access to spatial data.

The British Geological Survey developed a National Geosciences Index (Adlam *et al.* 1988) that would contain information about the sources, accessibility and nature of data held in both centralized and widely-distributed collections. It should be available to users in a form best suited to their requirements. The system allowed users to search for required information by geographical location; to provide only sufficient information for the enquirer to assess whether a particular type of data exists for an area, whether they are relevant to his needs and how to gain access to that data directly. The objective again, was not to create a large central data bank, but to increase knowledge of, and accessibility to the dispersed holdings of data.

Although GIS provides researchers and managers the automated approach to manage and integrate spatial data, it must be emphasized that computer technology is only as good as the source material. An important decision that must be made in the initial stages of a GIS project is on the primary data sets to be incorporated into the GIS. It was found that either a spatial index or a bibliography of spatial data were the top-ranked data sets. This emphasized the fact that public access to maps and spatial information depends on a link between bibliographic records, spatially indexed maps, and online vector/raster data. The menu system provided an indexing and

geographical search system for maps, photographic products, and imagery referenced by geographical location, and allowed interactive searches specified in a variety of ways from geographical windows, co-ordinates, or names.

A slightly different approach was followed by the South Australian Land Information System (Porter and Lores 1990) which was developed on a nodal database model (see Figure 2.3). Its underlying principle is to keep each separate database component as individual nodes of the model. Responsibility for the development, creation, data integrity and data currency of each node rests with the agency which "owns" the data. The duplication of data collection, storage and maintenance is avoided under this system by agreement being first reached on who shall be the data custodian.

The nodal approach is seen as an effective balance of the centralised/decentralised concepts and the most simple, flexible, practical and cost-effective method of achieving an integrated system. It allows each functional node to be equipped with hardware and software specific to its own needs. In the alternative hub model, all data is fed to, or accessed through, a central repository (the hub) for dissemination to the user community. (See Figure 2.2 page 30.) As a result the system is more complex, and technical knowledge of the data and its appropriate use is more likely to be geographically separated from the public access area.

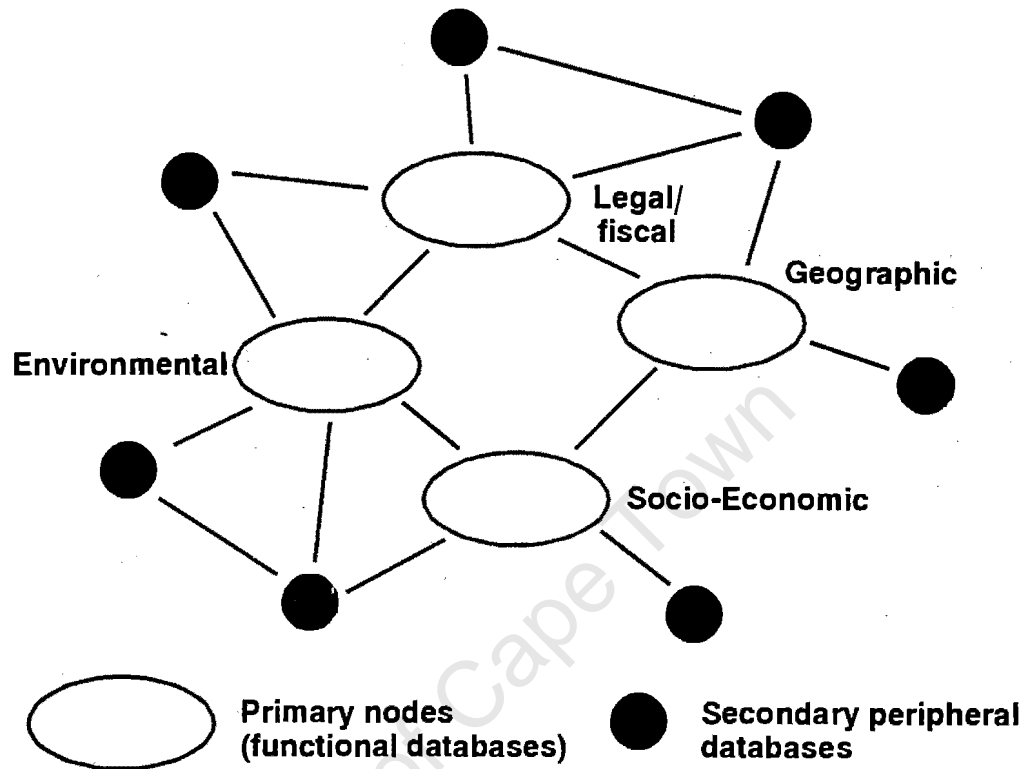


Fig. 2.3 *Nodal approach to a land database configuration (after Porter & Lores 1990)*

The major database nodes are:

- legal/fiscal - legal ownership and tenure system;
- geographical - digital cadastral data, survey and topographic data;
- environmental;
- peripheral databases such as facilities management and engineering data.

An advantage in South Australia was that the country was well-surveyed and mapped and that the boundaries of most land parcels were set out and recorded. This provided a sound graphical map series on which to base a digital cadastral mapping system. Durgin (1993) also pointed out the importance of a GIS base map tied to a good geodetic control system with a cadastral overlay.

Walker *et al.* (1992) distinguished between four types of data directories: those holding information about data sets held on one workstation; those holding information on data sets held within a single organisation; those that hold information about data sets available for a particular country; and those that hold information about data sets that apply to different countries, continents and the world generally.

#### 2.4.5 Metadata

Metadata is essential to a system that provides an index or catalogue to spatial data, because it provides the background information about the data so that the user can decide whether the data will be of value to him, and how to obtain access to the data. A metadata system was developed by the Midlands Regional Research Laboratory (MRRL) (Walker *et al.* 1992) to supply the user with information about where potentially useful data are held, by whom and in what form, together with details of restrictions on access and of payment required.

The central component of the structure is the metadata base, which contains information about potentially accessible data sets. This information comprises descriptions of each of the various data sets, including information about their content, structure and accessibility, and descriptions of the interrelationships of the data within and between data sets. The description of each data set also contains a specification of the spatial coverage to which the information in the data set applies and the period for which this information is relevant. The system makes provision for recognising words or concepts so that a number of data set descriptions are extracted from the metadata base. To be able to make a spatial inquiry, the area of interest is compared to the area covered by the data in the data set. If it overlaps the data set may be relevant.

The benefits of providing metadata (background information on the nature of data) were described by Cassettari (1993):

- prevention of duplication in the collection and storage of data;
- improved access to data;
- standardisation of data formats;
- increased flexibility in using data; and
- resultant enhanced value of data resources.

The type of information that may be included in metadata are:

- title, author and date of creation;
- sources of spatial information;
- accuracy, scale, quality and currency statements for sources;
- information about projection parameters and reference system;
- explanations of how source data were used to compile the data set;
- definitions of attributes and the rules by which they have been coded;
- rules and procedures adopted for data capture;
- results from geometric accuracy tests;
- types of analysis procedures performed on the original data and the constraints and quality of the results; and
- rules for the display and cartographic representation of data.

Medyckyj-Scott (1991) stated that the most common types of spatial meta-information systems (SMIS) are the inter-organisational ones. These are systems that contain information about data sets that exist in a particular geographical area, region, country or group of countries, and are held by a variety of organisations, e.g. NASA and World Climate Data.

#### **2.4.6 Temporal Database**

An often neglected, but very important type is historical and temporal data. Vrana (1989) identified the usefulness of temporal data in describing a land supply information system.

Such a system:

- is organised around the land parcel as its basic unit;
- contains information about the existing and forecasted land supply within a region;
- is accessible to public and private sector users; and
- is designed for monitoring changes in land supply inventory in order to assist in management and regulation of land development and to facilitate analysis of land markets.

Temporal data are transformed into historical information in the process of identifying the status of a parcel in a longer cycle of development, as well as in forecasting trends in the supply and demand for land development.

Historical and spatial information is used to manage land. A land information system (LIS) usually indicates the owner of a property at any particular time. When it changes ownership the name of the owner is normally replaced by the new owner, and it should be possible to track past ownership of a property. Also, changes of boundaries and land-use changes should be recorded to assist planning. Fully incorporating such historical information into a system, involves making use of the temporal nature already implicit in a LIS, as well as developing procedures for explicitly linking states of attributes at one time with those of preceding or subsequent times.



Three common aspects of spatial data handling affect historical information in a LIS. These are: (1) connecting a date as an attribute of a feature stored in a database; (2) creating files that store a series of events which can be used to reconstruct a "chain of events"; and (3) a method of updating an item of data but storing both the previous and updated item with a version connected to it.

Al-Taha and Frank (1991) said that a temporal GIS is required to analyse and understand phenomena by tracking their change over time, to predict future systems' behaviour by looking into their past, and to plan and take action that will lead to a desired future situation. This is used by politicians, meteorologists, planners and engineers. Questions that a temporal database should be able to answer are: When did John buy erf 61, or who were the owners of the houses north of Church street on 15 June 1985?

## 2.5 MAP LIBRARY

Traditionally a map library holds cartographic material such as globes, atlases and printed maps. In recent years it may also include videotapes, compact disks and digital data accessible by computers outside the library. Therefore Lai and Gillies (1991) proposed the term *spatial data library*. Card catalogues, indexes, guides and bibliographies are the traditional tools spatial data librarians have relied upon in the past. These techniques and approaches are, however, inadequate to meet modern

demands for information. The librarian should be able to use various spatial operators for extracting information from the database. Map-related queries may be undertaken either graphically or by using a logical statement.

Traditional catalogues do not make provision for (i) cross-referencing with non-bibliographic data; (ii) searching which involves arithmetic operations e.g. co-ordinate or window searching; or (iii) interfacing with non-bibliographic computer packages to allow, for instance, automated plotting of graphic indexes. The CARTO-NET system developed in Britain (Healey and Morris 1987 and Morris 1986) used a relational database and mapping software to catalogue cartographic materials. It provided (i) full bibliographic control of all aspects of map cataloguing; (ii) automation of graphic index production for maps at different scales and for aerial photography; (iii) linkages between data on aerial photographic coverage and the map catalogue; and (iv) linkages to gazetteer information/census information. It provided a computerized search method of the traditional card system and also showed a graphic index.

Holmes (1990) pointed out that such a system is distinctly different from a GIS. The primary difference is that the library is concerned with providing access to geographical information available in all types of media, essentially via a catalogue. As such, the catalogue is a superset of the GIS's purpose, which is to display and manipulate digital and geographical information for a specific site. The fact that a catalogue system may allow one to select and access the data, and then manipulate it, using graphics, mapping, image processing or GIS software, is simply an aspect

of its modularity. The manipulations are subordinate to the catalogue's basic access function.

One of the primary functions of a spatial data library is to retrieve geographical information in a response to a user's articulated needs for accessing and analysing data. There must be general information utilities for determining the availability and location of spatial databases, and the tools and methods to extract the spatial and non-spatial data. Lai and Gillies (1991) described a conceptual model of a spatial data library where the user not only has access to a bibliography of spatial data in the form of maps and atlases, but actually have direct access to digital databases. By using the sophisticated technology of a GIS he would be able to retrieve and manipulate spatial geographical data. This truly is a "geographic information system" in the library, which then fulfils its function, which is to provide all types of information to the community.

Mawdsley (1992) and Pearson and Sprunt (1986) mention the advantages of a computerized map library cataloguing system as portable, easy to use, giving faster access to library items and that its introduction is assisted by the availability of cheap software and microcomputers. A disadvantage is the time-consuming input of records into the computer.

All the papers pointed out the differences between cataloguing books, where author and title are important, and maps, where the spatial area is the most important. This could be described in terms of numeric methods (geographical co-ordinates), area

classification such as a postal code, textual methods (city name), or graphical methods where the area is indicated on the screen. Names, however, are not unambiguous because of duplication and different spelling or languages, and therefore geographical co-ordinates are preferred.

The terms "digital mapping", or "maps on screens" are often used with little attempt by the map library community to understand the great variety of materials of very different kinds subsumed within these broad definitions. The storage medium is important because it influences how much data can be accessed, whether that data can be altered, how the data is organised and what hardware and software configurations are required to access the information.

Perkins (1994) discussed the problems in accessing and using digital maps. It is not only the difference between raster and vector, but also the variety of file structures such as ASCII or DXF format which makes compatibility difficult. These are also accessed by a variety of different hardware platforms with different operating systems. Some of the software packages that may be available in a spatial data library are route planners, simple "paint" packages, map creation programmes, CAD systems and GIS. At the low end of the software hierarchy little user-support is needed, whereas at the high end users need considerable help to exploit the value of the product.

In addition to this complexity of media, format and software systems comes the inevitable dynamism of data. Whereas paper copy maps 200 years old can still be

accessed today, we have no way of knowing what systems will be operating as vehicles for accessing data even a decade in the future. The hard copy topographic map was readily available to all, as an artefact, whereas its digital equivalent might be fluid information only available to those with access to the necessary knowledge, software and hardware.

The elements that are listed in a bibliographic description are quite different in hard copy than in a digital format. For instance, accuracy-tagging of large scale digital products and date-stamping of different elements of the database and geo-coding coverage, are fundamental to successful retrieval of items to satisfy search criteria. The user will need to know the source of digitising and its accuracy (rather than the scale) since display scale is more often a function of the software being used to manipulate the data. If metadata is to be used for successful retrieval of digital products it is important that the data is collected in such a way that retrieval needs can be satisfied. For instance co-ordinate data would have to be collected if it was intended to offer graphical as well as textual area access to data sets.

Despite the proliferation of digital cartographic data sets, very few map libraries have yet decided to take on a dual role as a cartographic laboratory and conventional map library. The function of most map libraries remains to provide their users with access to hard copy published mapping, rather than offering the flexibility of digital data use to their customers. The digital format makes true data sharing possible, so it may therefore be more appropriate not to acquire digital data sets, but rather to access them when required over a network, to acquire collaboratively and share resources.

Data sharing, co-operative purchase and remote access may be the only viable economic route for map libraries to follow if they wish to continue to allow access to current cartographic data. It may therefore not be a good thing for libraries to buy digital data.

## **CHAPTER 3**

# **INVESTIGATION ON AVAILABLE SPATIAL DATA**

### **3.1 METHOD OF INVESTIGATION**

In commerce and industry there are predominantly two groups involved with spatial data: the suppliers and the users. A large number will obviously fall in both categories. The first objective of this study was to determine who are the suppliers of spatial data, and what type or format does this data have, and on the other hand, who are the users of data and what type of data they require. Then an evaluation would be made whether the supply meets the demand, and what steps could be taken to improve this.

The obvious source of information on data suppliers and the type of data available include literature, such as technical journals and other publications as well as contact with the well-known organisations and professions in the business, such as the Directorate of Surveys and Land Information and surveyors. The problem with literature is that it may be out of date and incomplete and it was therefore decided to contact the suppliers and users personally, and to obtain the necessary information from them.

This was done by means of a questionnaire which was compiled in such a way as to determine the type of data used and/or supplied by the person completing the questionnaire, or his organisation. Appendix A shows a copy of the questionnaire. Because this would be obtained from many different persons with varying degrees of knowledge of spatial data, it had to be compiled in such a way that it could be understood by everybody. It also had to cater for both suppliers and users of data at the same time, because many people could not be identified beforehand, as belonging to a particular group.

It is a well-known fact that only a small percentage of questionnaires sent to the public, are returned. Therefore steps had to be taken to ensure a large return. Personal interviews were conducted with as many people as possible, although for financial reasons this could only be done with people in the Western Cape. It also had the added advantage that much more and accurate information could be obtained in this way. It also provided further leads to other suppliers and users of data. To further improve the return rate of questionnaires that were posted, many of the people



were also telephonically contacted to explain the purpose of the investigation and to obtain their support.

For a descriptive survey the population of the survey must be carefully chosen and clearly defined, and precautions must be taken to safeguard the data from bias. The population of this survey were all the individuals, professions and organisations involved with spatial data. Within these parameters the population was taken as wide as possible so that the whole spectrum could be covered. It was further decided that the quantity of the population was not so important as the necessity to include all the different role players and professions. Appendix B provides a list of the groups that took part in the survey.

Professions and organisations were grouped in the following four categories (a detailed list is given in Figure 3.1):

- *Spatial data managers* - they are the custodians of spatial data and their main functions are the collection, processing and dissemination of spatial data.
- *Planners and construction* - use the base data supplied by spatial data managers for design work and then produce spatial data for use by themselves and allied professions.
- *Land use and development* - are concerned with the economic use and

development of land and the conservation of natural resources.

- *Services and information* - these are organisations providing a service or information to the public.

Spatial data managers	Air survey companies Mapping organisations (state and private) Cartographers Surveyors Major GIS users and vendors Local authorities SANDF
Planners and construction	Civil engineers Town planners Architects Landscape architects Quantity surveyors Draughtsmen
Land use and development	Surveyor-General Deeds office Nature conservation Ecologists Geologists Botanists Agricultural research Farmers Property developers Estate agents Valuers
Services and information	Libraries Telkom Traffic police Economic analysts CSIR Education Meteorological office

Fig. 3.1 *Categories of spatial data suppliers and users*

### 3.2 SPATIAL DATA SOURCES

South Africa has since the beginning of the 20th century worked on a survey system that has ensured high accuracy of all topographical and cadastral data. This system is based on beacons that are spread across the country and allows for a single geographical referencing system. The Land Survey Act (Act 9 of 1927) makes it compulsory for all cadastral surveys to be based on the national co-ordinate system, which makes it possible to relate all land parcels to each other. This also ensures that there are very few disputes about cadastral boundaries in South Africa. The state is responsible for establishing and maintaining the necessary trigonometrical beacons with co-ordinates and heights based on the national co-ordinate system across the country. This has ensured that not only cadastral surveys, but also most topographical and engineering surveys are based on the national system. This has made it possible to relate most survey data to each other, although they may have been done for different reasons, at different accuracies or different times.

The national control survey network is currently being readjusted and converted onto the WGS84 (World Geodetic System 1984) reference framework. This is being done because new values have been determined for the reference ellipsoid, using GPS (global positioning system) observations. This will, however, not result in erroneous co-ordinates, as long as the necessary conversions between the systems are applied.

The spatial data managers are the largest suppliers of spatial data. Almost 60% of all

the respondents in the investigation reported that they supply data in digital format, usually vector maps or plans and/or co-ordinates. Figure 3.2 shows that 42% of all respondents supply topographical maps and 40% supply city maps. The latter include land use, zoning, engineering and building plans. Spatial data managers supply mainly topographic data (62% of respondents) to other users, and when internal use is included this figure increases to 76%. This can be supplied by almost all (95%) in digital format.

	All respondents	Data managers	Planning & Construction
Digital data	57%	76% (95%)	63%
Topographical	42%	62% (76%)	50% (75%)
City	40%	52% (62%)	63%
Political & Admin.	34%	52% (57%)	25%

Fig. 3.2 Major types of spatial data supplied. Percentages between brackets are when supply for own use are included.

Organisations and individuals in planning and construction are also major suppliers of data. However, this is often after they have received the base data and enhanced this by adding information, such as a town or building design.

### **3.2.1 State**

#### **3.2.1.1 Surveys and Land Information**

The Directorate of Surveys and Land Information has been established in terms of the Land Survey Act (Act 9 of 1927), which also makes provision for the appointment of the Chief Surveyor-General, and sets out their functions:

"The Chief Surveyor-General shall be in charge of such geodetic, topographic and cadastral surveying and land information services as the Minister may direct, and shall -

- (a) promote and control all matters affecting such surveys and services;
- (b) supervise and control the survey and charting of land for purposes of registration in a deeds registry;
- (c) conduct such trigonometrical, topographical, and other relevant survey operations as may be required;
- (d) prepare, compile and amend from time to time, as the circumstances necessitate, such maps and other documents as may be required; . . ."

## **1. National Land Information System**

The National Land Information System (NLIS) is a computerised management system that has been established to provide a meaningful, co-ordinated and integrated land information system on a national level. It will enable planners and other decision-makers to have timeous access to information on which to base decisions on the physical, socio-economic, scientific and security development of the country. The major functions of the NLIS are to maintain a register of suppliers and users, maintain an index of data, allow on-line inquiries to the index and to ensure data transfer.

Recently the objectives for the NLIS have been scaled down and it is now envisaged that a metadata base will be provided to users on CD-ROM (Clarke 1995).

A National Exchange Standard has been produced to meet the need to exchange (transfer) geo-referenced information in computer-compatible form from one database to another, either within the same computer system or between different computer systems. A format is recommended whereby information in either vector or raster format (or a combination of both) or alphanumeric data only, can be exchanged.

## **2. Map series**

As part of its primary function the department publishes and maintains the following map series which cover the whole country:

1 : 50 000 Topographical maps

1 : 250 000 Topographical and topo-cadastral maps

1 : 500 000 Topographical and administrative maps and aeronautical charts

1 : 1 000 000 World aeronautical (ICAO) chart

1 : 2 500 000 Map of Southern Africa

1 : 7 500 000 Map of Africa South of the Sahara

The department is currently digitising these maps.

It also publishes a 1 : 10 000 orthophoto map series covering all the major centres and areas of development. Aerial photographs at various scales are available for the entire country.

## **3. Co-ordinates**

Co-ordinates and heights are available in paper and digital format for all the trigonometrical beacons and town survey

marks as well as triangulation plans showing their positions. Heights of bench marks are provided along with maps showing their positions.

#### **3.2.1.2 Surveyor-General**

The Land Survey Act has allocated the offices of the Surveyor-General the responsibility to approve, preserve and make available to the public all diagrams, general plans and other cadastral documents related to all properties in his province. All real rights related to land, such as ownership and servitudes, are unambiguously defined in the numerical cadastre. This information is available in hard copy format and is also now being captured onto a digital cadastral information system. This will be linked to the database of the Deeds Office, which provides information on ownership and other real rights.

#### **3.2.1.3 Department of Water Affairs and Forestry**

The Department of Water Affairs and Forestry has the responsibility to manage the water resources of South Africa on a national level. It therefore produces maps in hard copy and digital format of all the major rivers, dams, catchment and



drainage regions in the country. It also operates one of the largest and most successful GISs in South Africa. A large variety of data such as vegetation, statistics, geological and topographical is put on the GIS.

#### **3.2.1.4 Department of Agriculture**

Soil type maps and grazing capacity for agricultural purposes are produced, mainly for farmers. A GIS is used to determine the agricultural potential of land. The Agricultural Research Council uses a GIS for meteorological research and produces meteorological maps and data in hard copy and digital format.

#### **3.2.1.5 South African National Defence Force (SANDF)**

The SANDF has several units responsible for surveying and mapping and produces a variety of maps in both hard copy and digital format.

#### **3.2.1.6 Central Statistical Services**

This department is the official collector and supplier of statistical information, including the population census, and is responsible for demographic data. Statistics can be provided

for geographical areas and with a 7-digit code it distinguishes between statistical region, magisterial/census districts and urban areas.

### **3.2.2 Statutory Bodies and Utilities**

#### **3.2.2.1 ESKOM**

To fulfil its function of providing electricity to the country, ESKOM has its own surveyors and engineers using a wide variety of data to provide mainly topographical and geological maps and co-ordinates. This is used for environmental impact studies, route selection and the acquisition of servitudes.

#### **3.2.2.2 TELKOM**

TELKOM has established a GIS for the urban areas containing all the cadastral data, their communication lines and a digital elevation model (DEM).

#### **3.2.2.3 TRANSNET**

The major suppliers of spatial data in TRANSNET are land surveyors and engineers responsible for information on, and

management of their properties. Most of this data is produced for internal use only.

#### **3.2.2.4 Council for Geoscience**

Geological maps in hard copy and digital format at various scales accompanied by explanations are published for certain areas of South Africa. The country is fully covered by the 1:250000 series using the maps from the Directorate of Surveys and Land Information as a base. A large number of geologically related publications is also available.

#### **3.2.2.5 Local Authorities**

Most local authorities have their own surveyors, engineers and town planners that are responsible for the provision of maps, co-ordinates and heights for the provision and maintenance of services in their areas, and details of urban land use and buildings. This may vary from small manual productions to large automated systems and GISs. This may include topographical, engineering, cadastral and all the types of city maps as well as aerial photographs and co-ordinates related thereto.

### **3.2.3 Private Organisations**

#### **3.2.3.1 Mapping**

Some organisations perform both aerial surveys and mapping, while others specialise in mapping. A wide selection of products is supplied, including topographical, engineering, cadastral, land use, street maps, atlases and orthophoto maps in hard copy and digital format at various scales. Aerial photographs, diapositives and remote sensing images are also produced. The aerial survey companies mostly produce data for specific clients according to their requirements.

#### **3.2.3.2 Surveyors**

Land surveyors are responsible for cadastral surveys and are therefore responsible for the production of diagrams, general plans and sectional title plans. They also have to mark all boundary corners with beacons and determine co-ordinates of these.

Surveyors are generally involved with topographical and engineering surveys and the production of maps and plans for land development and services such as roads and dams, and

setting out these projects.

#### **3.2.3.3 Engineers**

In the design, construction and maintenance of services such as roads, dams and pipelines, engineers will produce design and construction plans. Most of this design work is done on computer and therefore available in digital format.

#### **3.2.3.4 Town Planners**

Town planners make use of topographical maps for their design and development work and produce township layouts, structure plans, zoning maps and site plans, usually in paper and digital format.

#### **3.2.3.5 Architects**

Digital contour maps are used for the production of building plans, coverage and bulk plans, most of these also show heights in various forms.

### 3.3 USERS OF SPATIAL DATA

The biggest demand for spatial data is by individuals and organisations in the planning and construction group. Figure 3.3 show the number of respondents that uses a particular type of data. Political data includes provincial, magisterial and municipal boundaries. City data includes land use, zoning, engineering and building plans which is mostly used by town planners, architects and engineers. This data may be in digital or hard copy format.

	All respondents	Data managers	Planning & Construct.	Land use
Topographic	75 %	81 %	88 %	65 %
Political	75 %	76 %	63 %	76 %
Cadastral	68 %	76 %	75 %	65 %
Road	66 %	76 %	75 %	53 %
City	64 %	67 %	88 %	53 %

Fig. 3.3 Demand for different types of spatial data

An indication of the availability and use of digital data is given in Figure 3.4. It is widely used by individuals and organisations from spatial data managers and planning and construction, but to a much lesser extent in the land use and information and service fields. This is because special software is required to manipulate spatial data which is used frequently by the first two groups but not by the last. It must therefore

be realised that the latter may not have the ability or the need to access or use digital databases. In all the groups the suppliers and users are generally the same individuals or organisations.

	All respondents	Data managers	Planning Constr.	Land use	Info & Services
Supply	64%	95%	63%	35%	43%
Used by	55%	71%	63%	35%	43%

Fig. 3.4 *Number of respondents that supply or use digital data.*

51% of the respondents reported on the National Exchange Standard, and of these only 33% complied with the standard. The responses were varied: some replied that there was no need for it, while others felt that it was actually ahead of its time and that South Africa was not yet ready for it. There was a general lack of knowledge of its purpose and use, and some education in this field is clearly necessary.

### 3.3.1 Spatial Data Managers

This group includes mapping organisations, surveyors, cartographers, GIS organisations and other organisations directly involved with spatial data. In general they are both users and producers of spatial data. Many organisations match the requirements of these practitioners because the market for these kinds of data is large, the application areas are reasonably

well understood and are supported by legal frameworks. They know exactly what they need and where to find it, because they are the specialists. If, what they need is not available, they produce or manufacture it themselves. Some of the advantages that they will derive from a more effective supply and distribution of data are:

- less duplication of data and a better utilisation of resources;
- cheaper data and therefore a reduction in the cost of information; and
- new uses and additional markets would be found for the data.

Spatial data managers use mostly topographic, cadastral, political/administrative and road data, which is often in digital form. Much of this is enhanced by adding value to it and then supplied as information to other users. Examples are air survey companies and mapping organisations using aerial photographs for map production, land surveyors producing cadastral diagrams and GIS organisations doing spatial analysis.

### **3.3.2 Planners and Construction**

GIS is increasingly being used for spatial planning. This group has infused its tradition and methods of conceptualisation of the spatial world into the GIS. Some commercially available systems cannot disguise their roots in a planning tradition which used transparent maps overlaid one on the other to find areas of certain characteristics. They have developed from using



traditional methods, then CAD systems and now the analysis power of GIS.

Planners, architects and engineers use mostly topographical, cadastral and zoning data to produce city maps such as layout maps and design plans for construction. Most of this data is in digital format and is supplied to the next user in this form.

### **3.3.3 Land Use and Development**

These professions are concerned with land use for agriculture and the provision of housing and environmental conservation. They use a large variety of data e.g. topographical, cadastral, vegetation, soil, geological, and meteorological. This data is not static, e.g. vegetation changes, endangered animals become extinct and housing changes continuously. Therefore regular updates are essential, and temporal information may be very valuable. Spatial data is often compiled and supplied for specific users and specific purposes, e.g. geological maps or grazing capacity for agriculture.

Environmental scientists such as those charged with inventory, modelling and monitoring physical processes in the earth's atmosphere use spatial information which is difficult to capture. Instead of creating databases that are complete models of the world they have to make do with samples, with approximations. They deal with interpolation and extrapolation and must

trade certainty for probability.

Although it has been indicated in various articles (e.g. Grant 1993, Hiland *et al.* 1992 and Wilkinson 1995) that aerial photographs and satellite images are important sources of data, very few respondents have indicated that they actually use them (38% and 25% respectively). Even orthophoto maps are used by only 56% of the respondents, as compared with 100% of planners. This may be due to a lack of knowledge of the data, or because it is not readily accessible.

#### **3.3.4 Services and Information**

These professions are responsible for providing services such as TELKOM and traffic police, and for providing information such as universities, CSIR and libraries. The latter group normally amend or enhance the data before passing it on to the final user, while the former group is usually the final user.

#### **3.3.5 Politicians and Entrepreneurs**

Decisions regarding space are made by politicians and entrepreneurs. They decide where a new shopping complex or factory is to be built or how agricultural land is to be developed. They use advice from the groups listed above and they make use of output from GIS and other spatial analysis tools

when it has been duly condensed and prepared such that the information can be understood and combined with other information to be considered (mostly legal and economical) and is included in formalised and legalised decision processes. Demographic, cadastral, zoning information and census data are used in the decision making process.

### 3.3.6 General Public

Last, there is the public at large, which lives in the spatial world. Enormous amounts of spatial information are constantly being processed, mostly in intuitive and informal ways, before being stored. Decisions are often made in 'real time'. All analyses and use of spatial data is done with the general public in mind, and is supposed to be for their benefit. Figure 3.5 shows the spatial data supplied for the use and/or benefit of the public.

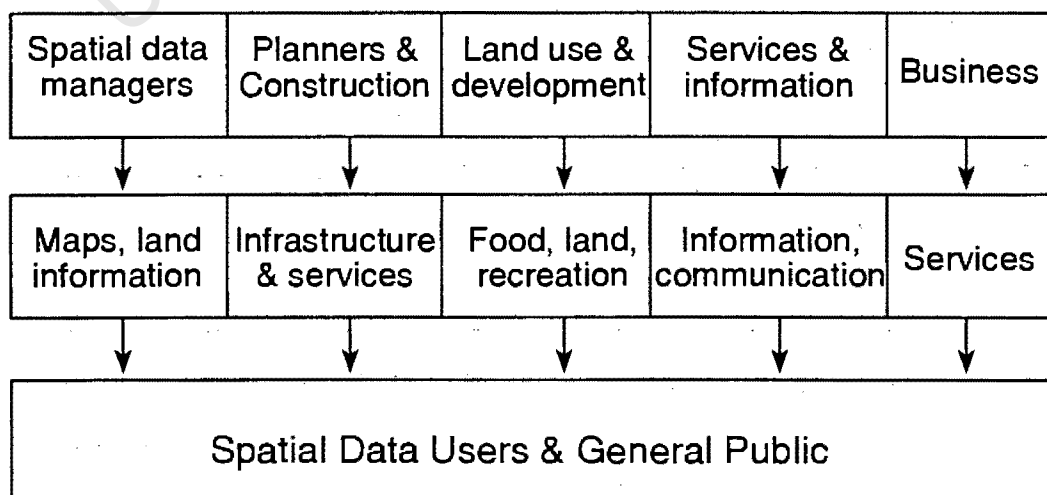


Fig. 3.5 Types of spatial data supplied to the users

### **3.4 CONCLUSIONS ON THE AVAILABILITY OF SPATIAL DATA**

The responses of the various groups were somewhat different. The spatial data managers identified very few shortcomings in the supply of spatial data. As a few respondents put it: "If we need something that does not exist, then we make it". It is the function of this group to supply spatial data, and they are therefore well-informed about the existence of data and where to obtain it.

The planners and construction group were reasonably well-informed about spatial data sources and most of them used the latest technology. Some respondents however reported that orthophoto maps and satellite imagery may be useful.

There is a need for current land use data for urban, but especially for rural and agricultural areas. This is required by the agricultural sector and ecologists. More complete coverage of soil types are also required. Although this is supplied by the Department of Agriculture, they do not have complete coverage of the country. There is a general need for vegetation data, which is supplied to a limited extent by the departments of Water Affairs and Forestry, Agriculture and Surveys and Land Information. None of these however provide sufficient data for the demand.

A number of users has the need to link an erf number with a street address and ownership. There should be some link between the databases of the offices of the Surveyor-General, Registrar of Deeds and the local authority. Although it is an objective of the cadastral information system, problems are encountered because different databases are used; the Deeds Office uses a traditional alphanumeric database with no spatial abilities.

An indication of the demand for a central information system or a metadata base is the fact that a number of respondents said that historical data would be useful, but it is not available. The South African Library and the National Archives have some historical data, but it is not well catalogued and the possible users are not aware of this. It is not worthwhile for a user to conduct long searches which may prove fruitless. If information on available data was accessible in a metadata base the user would be able to ascertain whether any suitable data exists.

A central information system must therefore be established that could serve as a communication link between data suppliers and users. The suppliers must provide this information system either with the data itself, or sufficient background information about the data, in other words, metadata. It is essential that a complete record of data is maintained over the total spectrum of spatial information. The information system may be computerized or manual, although to be effective it should be a computerized database system. All possible users of spatial data must have access to the information system and be aware of its existence. It should make provision for all categories of users, from the person that is looking for general information, to the specialist who

is conducting research and requires specialised information. Cassettari (1994) has advocated the idea of a "spatial data hypermarket" where users can view and purchase spatial data from a wide range of suppliers. This would promote the supply of spatial data and increase its use. Research will therefore be directed towards the development of an information system or database to provide better access to spatial data.

## **CHAPTER 4**

# **DESIGN OF GIS DATABASE**

## **4.1 INTRODUCTION**

### **4.1.1 Data Management**

A database system is basically a computerised record-keeping system; i.e. a computerised system whose overall purpose is to maintain data (or information) and to make that information available on demand. Data refer to the values actually stored in the database and information to the meaning of those values as understood by a user. The information concerned can be anything that is deemed to be of significance to the individual or organisation the system is intended to serve - anything, in other words, that is needed to assist in the general process of running the business of that

individual or organisation.

A database system involves four major components, namely data, hardware, software and users. The data in a database should be both integrated and shared. By "integrated" is meant that the database can be thought of as a unification of several distinct data files, with any redundancy among those files either wholly or partly eliminated. By "shared" is meant that individual pieces of data in the database can be shared among several different users, in the sense that each of those users can have access to the same piece of data (and different users can use it for different purposes). The database should be independent of its applications and its design should concentrate upon the data that is actually collected and stored. The hardware consists of the secondary storage, the processors and the associated main memory.

Between the physical database itself and the users of the system is a layer of software, the database management system (DBMS). All requests from users for access to the database are handled by the DBMS, and it is thus a shielding of database users from hardware-level details. The DBMS is used for the storage, manipulation and retrieval of data from a database. The users are the application programmer, responsible for writing application programs that use the database, the database administrator and the end user. The end user can have access to the database via a database query language or menu-driven interfaces.



A database consists of some collection of constant data that is used by the application system of some organisation. Any distinguishable object that is to be represented in the database is called an 'entity', e.g. a person or an erf. It is important to note that there are 'relationships' linking those basic entities, e.g. a person owns an erf. These relationships between entities can be utilized by the DBMS for manipulation and retrieval purposes. Entities also have properties, e.g. the area of an erf or the age of a person.

The advantages of having centralised control via a database are:

- redundancy can be reduced;
- inconsistency can be avoided;
- the data can be shared;
- standards can be enforced;
- security restrictions can be applied;
- integrity (accuracy of data) can be maintained; and
- conflicting requirements can be balanced.

Underlying the need for better geographical information management is its burgeoning volume. Having even a few digital geographical files without a suitable index is difficult enough to manage, but the earth resources satellites will inundate existing facilities unless provisions are made for sorting and indexing the huge volumes of information which can be expected. Even the volume of traditional media is growing rapidly. As new

editions of maps, aerial photographs and environmental data are accumulated, it becomes increasingly difficult to organise, and for the user to find these materials.

A second driving force for better management is the increasing need for integrated geographical information. Environmental impact studies require historical and ecological analyses of regions. Development plans need consideration of natural and cultural factors and modern scientific research frequently requires a more integrated approach, looking at the distributions of various environmental components, such as soil type or an insect species.

A third influence is a recognition that when manipulated by a computer, geographical information can be remarkably informative. Instantaneous graphics and maps, statistical analyses, coordinated maps and images can be used together to quickly provide a substantial understanding of a place.

#### **4.1.2 Philosophy of the Database Design: Affordability and Simplicity**

The purpose of this research was not to develop a custom-made database for this application, but to use a standard commercial GIS to provide the required information and facilities. It is clear that a new system developed specifically for this purpose would be more effective. However, such a system would be very expensive to develop and expensive to set up and maintain. Furthermore a small- or medium-sized local authority would not

have the expertise or the funds to develop and maintain it.

Because many local authorities already have GISs operating, it would be relatively simple and affordable to add this facility to it. The purpose was rather to use existing hardware and software to provide the necessary services and information.

## **4.2 FORMAT OF DATA**

### **4.2.1 Extensive Geographic Information System**

An extensive GIS database would reflect the total needs for spatial information handling. It would include a complete cartographic database, cadastral, deeds, land use, vegetation, physical resources, cultural and demographic data. It would have the following advantages:

- users would have direct and on line access to all data; and
- it would maximise the sharing of the most current and accurate data available and significantly reduce redundant data collection.

The disadvantages of a GIS with an extensive database are:

- a very large database would be required to store all the different data

sets comprising maps, photography, imagery and literature. Such a database would be cumbersome and difficult to manage and maintain;

- data obtained from different sources may not be compatible in the same GIS and database;
- data from different sources will have different accuracies and different data standards;
- a large volume of data is not yet available in digital form; and
- data producers are often unwilling to supply data both from reluctance to take responsibility for errors in the data, and because of concerns over pricing and copyright.

With increased availability of GIS applications software, more geographical information is processed each day. Managing this growing volume, so that access to spatial data is maintained, requires powerful computer processing and large computer databases that are linked over a network. Without this capability, the large quantity of spatial data is unmanageable and becomes disorganised and inefficiently accessed. Holmes (1990) and McBride *et al.* (1991) discussed the need for better access to and management of geographical information and outlines six reasons for networking: (1) the volume of geographical information is immense; (2) it allows integration of diverse geographical information; (3) more information is gained about an area of interest; (4) it provides an efficient method to manage and monitor human activities; (5) all information is shared equally; and (6) remote access to otherwise inaccessible geographical information is established.

#### 4.2.2 National Spatial Data Infrastructure

The purpose of a national spatial data infrastructure (NSDI) is to create an information highway consisting of databases and communication networks that will enable spatial information to be more easily located, communicated and used (McLaughlin & Nichols 1994). Figure 4.1 illustrates the components of a NSDI. These can be described as:

- *Sources of spatial data:* the groups and individuals acquiring, creating and managing spatial data and databases in the private sector and government.
- *Databases and metadata:* sets of organised spatial data and information about those sets.
- *Data networks:* the communication highways in various forms (e.g. telephone lines and local area networks) linking databases, sources and users.
- *Technology:* the hardware and management of databases.
- *Institutional arrangements:* coordination of the organisations involved.
- *Policies and standards:* the data communication rules and issues such

as privacy and pricing.

- **Users:** the individuals and organisations who use the data and also add new information.

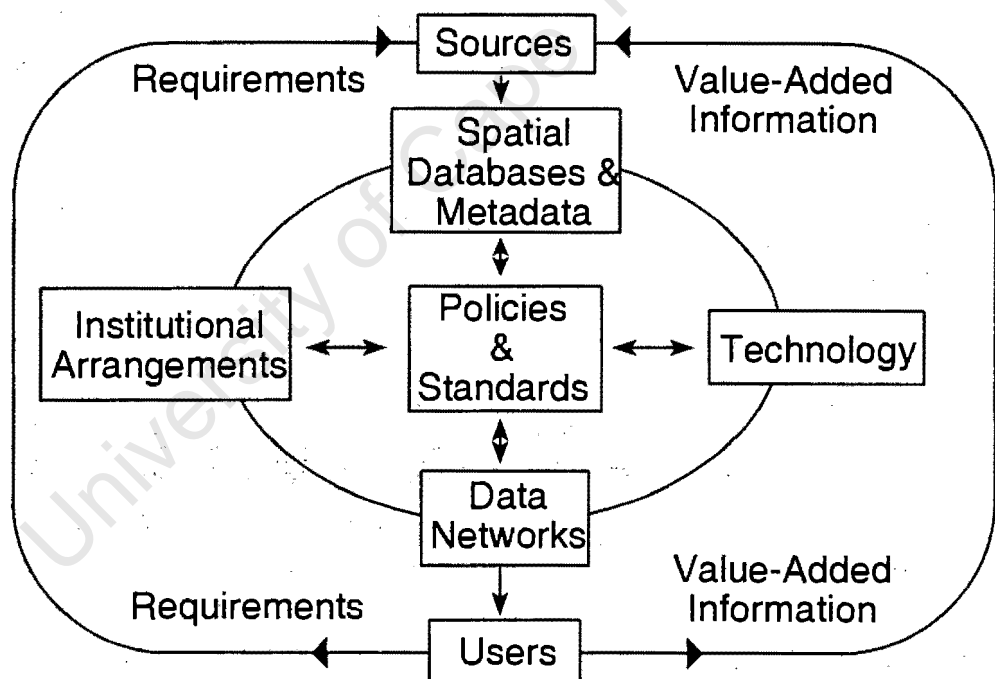


Fig. 4.1 *The components of a NSDI (after McLaughlin & Nichols 1994)*

The NSDI would create a national network linking databases and users and enhancing the accessibility, communication and use of geographically

referenced data. This will however be a very complex and expensive system to operate. It will be difficult to get the cooperation of all the organisations involved with spatial data, and many will not allow others to have access to their databases.

#### **4.2.3 Metadata Base**

A metadata base stores data about data, it provides background information on the nature of data, the sources from which it was derived and the overall quality of the material. This is valuable for those intending to use the data for deriving further information. Other information may include procedures of data capture, accuracy, types of analysis performed and rules for the cartographic display of the data. A metadata base provides "pointers" to other databases which may not be accessible and also to other forms of spatial information.

Some of the advantages of a metadata base are:

- it provides references to digital and non-digital data, e.g. maps, aerial photographs and literature;
- it will provide users with a complete index of available data and the format of this data;
- it should be easier to update because only the metadata would be updated and this is often of a more permanent form;

- the system would be relatively cheap because it can operate on a personal computer or a small network; and
- because it provides data about the data the user can decide whether the data would be suitable for his needs before acquiring or paying for it.

The disadvantages are:

- direct access to the actual databases are not provided. In a hybrid system this may however be done in certain cases; and
- the currency of information will be dependent on the suppliers; this may vary from one supplier to the next.

#### 4.2.4 Conclusions

An extensive GIS and a NSDI are ideal for individuals who require continuous access to the spatial data of a particular project and who make extensive use of this. Researchers who will work for a considerable period of time on a large project and who have the financial means and hardware will need direct access to the data. This is however not suitable for the general user who does not have the expertise, the finances or the hardware to operate such a system. For these users that may require any type of data, a metadata base would sufficiently meet their needs. It is also more flexible regarding the amount of detail that needs to be stored.



In Paragraph 3.3 it was shown that more than half of the data users in the Land use and Information and Services groups did not make use of digital data. A conventional GIS and NSDI would therefore not provide the service they need. They would, however, be able to make use of a metadata base which was situated at a central location, provided that it was easy to use and was able to direct them to the suppliers of the data they required. A metadata base running on a standard commercial GIS will be developed to provide sufficient information about spatial data to improve the economic supply and use of data.

Four different types of DBMS software may be used to implement the data model. These are

- inverted list systems;
- hierarchical systems;
- network systems;
- relational systems.

An inverted list is a table or index of data addresses that indicates all the data (records) that have common characteristics. An index is used to assist the retrieval of records. The hierarchical data model has a very efficient structural form when data relationships follow a purely nested one-to-many pattern. However, its limitations cause undesirable redundancy. The network data model represents data as sets of record types and relationships between pairs of record types. It is, however, complex and difficult to use. The

relational model has a more formal mathematical foundation than the others - both in the description of its structures and in the languages specified for it.

A relational DBMS is more suitable for many small and medium-sized database applications where flexibility and quick system development are important. Network and hierarchical systems are good for designing and implementing large databases with well-defined queries, transactions and applications.

The relational database system is normally used because of the following advantages:

- rigorous design methodology based on sound theoretical foundations;
- all the other database structures can be reduced to a set of relational tables, so they are the most general form of data representation;
- ease of use and implementation compared to other types of system;
- modifiability, which allows new tables and new rows of data within tables to be added without difficulty; and
- flexibility in ad hoc data retrieval because of the relational join

mechanism and powerful query language facilities.

## **4.3 DATABASE PLANNING**

### **4.3.1 Design Process**

Database design is the development of the structure of the database, as well as the definition of its contents and the specification of constraints to be placed on the validity of data. The goals of database design are: to assure that all data needed to satisfy users' requirements are stored in the database; to eliminate redundant data; to provide a way to understand the organisation of the data; and to support the specified processing requirements and performance objectives. (Elmasri and Navathe 1989, Fernández and Rusinkiewicz 1993).

Typically, the process of database design involves (see Figure 4.2): (a) the collection of data related to the intended uses of the database; (b) the interpretation of users' needs and the development of a conceptual schema; (c) the mapping of the conceptual schema into the data model of the database management system; (d) the specification of storage structures and paths; and (e) the implementation of the database.

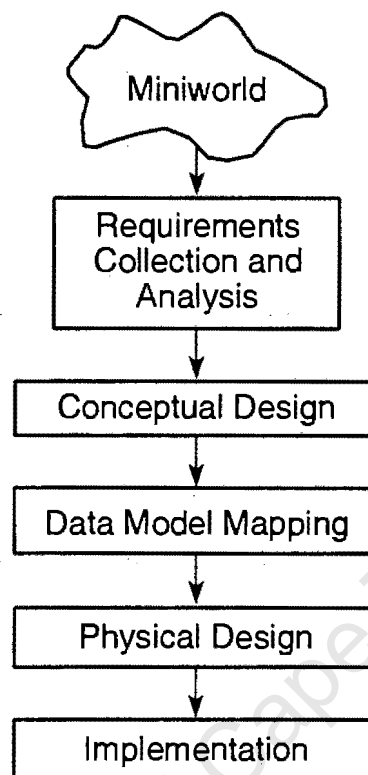


Fig. 4.2 *Typical steps in database design*

The principles for planning and designing of a database are the same for an alphanumeric database, a GIS database or a metadata base. The method discussed in the following paragraphs can therefore be applied to any type of database, although the examples are given for a metadata base.

#### **4.3.2 Analysis of Information Needs**

All users, irrespective of their particular specialisation, will be using data representing the topographic (and cultural) background to our everyday activities, and according to Haywood (1986) "How they use this data need

not, and should not, be the concern of the cartographer unless he has been commissioned to produce a special data set or database for a customer". Potential users of (digital) data must be made aware of this. In the same way as the user of a general topographic map had to create a system for effectively applying the information in that map to his own needs, he may have to change, reorganise, or add to the digital data to make it compatible with his software, hardware, and applications.

In this study it was found that the biggest need is to inform users of the spatial data available, in what form it is and where to obtain it. The purpose of the database is therefore to promote the business principle of supply and demand.

The objectives may be summarized as follows:

- keep a directory of all the spatial data in an area;
- provide the user with sufficient information on the data so that he may decide what data to obtain and the source and cost of this data; and
- ensure that the provision of data also benefits the suppliers.

### 4.3.3 Information Model

In setting up an information model we must answer three questions:

1. What are the important entities of this information system?
2. What are the associations (or relationships) between those entities?
3. How can we best portray the information architecture in graphic form?

An entity is a person, place, object, event or concept about which data must be recorded. The three most important entities are suppliers, users and spatial data. Spatial data can be broken down into maps, photographic images, co-ordinates and digital data. The relationships are shown in Figures 4.3 and 4.4 in the form of entity-relationship (E-R) diagrams. Each supplier supplies different types of spatial data to different users, therefore a one-to-many (1:m) relationship.

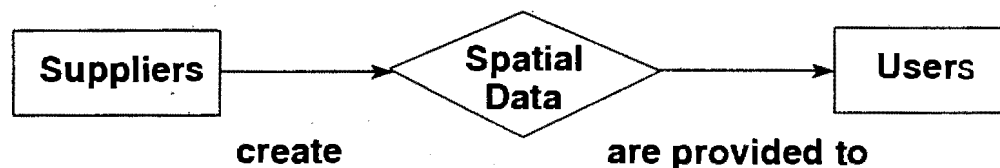


Fig. 4.3 Relationships between entities

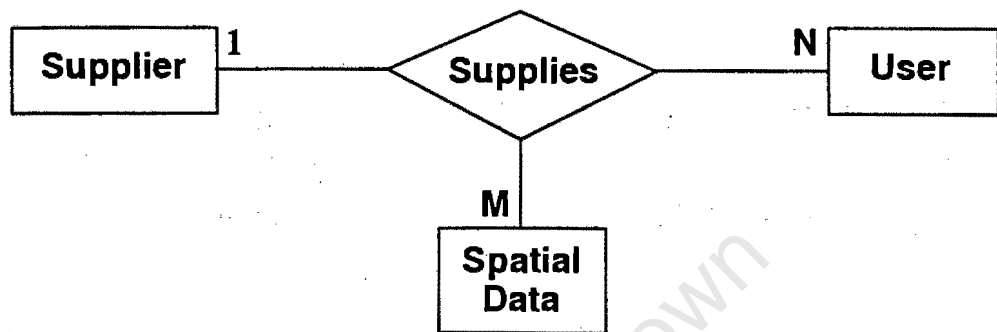


Fig. 4.4 *A supplier supplies many types of spatial data to many users.*

## 4.4 REQUIREMENTS ANALYSIS

### 4.4.1 Scope of the Database

In Paragraph 4.2 it was shown that three possibilities exist: an extensive GIS containing all the data, a national spatial data infrastructure or a metadata base. Considering the amount of data and the type of information required from it, a metadata base was found to be the most practical and viable solution. It would fulfil the needs of users to direct them towards the suppliers of the specific data they needed.

The system should be able to provide for the specialised needs of the professional user who may require detailed information, and also for the general user who will only use it occasionally. Access must therefore be user-friendly but also provide sufficient information for the regular user.

#### 4.4.2 User Views

This is the data required by a particular user to make a decision or carry out some action. The purpose of this is to determine the exact requirements of all the users of the database. A person requiring metadata concerning a map will, for example, need the data given in Figures 4.5 and 4.6.

Reference	Type of map	Title
Scale	Source	Supplier address
Date	Edition	Area of cover
Contour interval	Projection	Detail shown
Price		

Fig. 4.5 *Metadata required by map user*

Title	Scale	Area of cover
Source	Supplier address	
Date	Date of aerial photography	
Edition	Contour interval	Projection
Detail shown	Price	

Fig. 4.6 *Metadata required for an orthophoto map*

This must be determined for all the users of the database, whether they are users of spatial data or suppliers.



A data dictionary must be compiled for all the data items; these are attributes such as data type (e.g. alphanumeric), field length, range of data item (e.g. only positive values), or who is allowed to edit the contents.

## 4.5 CONCEPTUAL DESIGN

After the requirements analysis is completed each user view is modelled separately and the relations merged to form a conceptual schema.

### 4.5.1 Data Modelling

This is the process of identifying and structuring the relationships among data elements, and will result in a set of relations in "third normal form" (3NF) for each user view. An unnormalized relation is a relation that contains one or more repeating groups, e.g. the 'source' data in Figure 4.7. The purpose of normalization is to remove redundancies (or duplication) from the database and to reduce complex user views to a set of small, stable data structures. Redundancy may create errors when the database is updated, e.g. when records are added, deleted or edited this may have to be done in more than one place in the database. If all occurrences of a record are not edited the data would be inconsistent. This is illustrated in Figure 4.7 which shows part of the relation for the entity *Map* which has been formed from the metadata required for maps. The source, address and telephone columns

contain redundant data.

REFERENCE	TITLE	TYPE	DATE	ED	SCALE	CONTOUR_INT
3319CA	Bain's Kloof	Topographic	1979	2	1:50000	20 m
3319	Worcester		1980	3	1:250000	50 m
WEL_10	Wellington	Topographic	1985	1	1:500	2 m
3317	Cape Town	Admin	1982	1	1:500000	100 m

MAGIST_DIS	SOURCE	ADDRESS	TELEPHONE	PRICE
No	Surv. & L. Inf	P.Bag X10 Mowbray	6854070	R6,50
No	Surv. & L. Inf			
No	Wellingt.Mun.	Box 12 Wellington	31121	R15,00
Yes	Surv. & L. Inf	P.Bag X10 Mowbray	6854070	R6,50

Fig. 4.7 The unnormalized relation 'Map'

A relation is in "first normal form" (1NF) when it contains only elementary (or single) values at the intersection of each row and column. The relation *Map* is not in 1NF because the references 3319CA and 3319 have identical values for the attributes SOURCE, ADDRESS and TELEPHONE. A shorthand notation for the relation is as follows:

MAP (REFERENCE, TITLE, TYPE, DATE, ED, SCALE, CONTOUR\_INT, MAGIST\_DIS, SOURCE, ADDRESS, TELEPHONE, PRICE)

where REFERENCE is the "primary key". A primary key is an attribute or data item that uniquely identifies a record in a relation. This relation could be converted into 1NF by splitting it into two relations:

MAP (REFERENCE, TITLE, TYPE, DATE, ED, SCALE,  
CONTOUR\_INT, MAGIST\_DIS, SOURCE)

MAP\_SOURCE (SOURCE, ADDRESS, TELEPHONE, PRICE)

The "normalized relation" *Map* is shown in Figure 4.8. ED is dependent on both REFERENCE and DATE, which is known as "partial dependency", i.e. dependency on the key attribute (REFERENCE) and a non-key (DATE). For a relation to be in "second normal form" (2NF) there must be no attributes partially dependent on the primary key, as in the case of ED.

REFERENCE	TITLE	TYPE	DATE	ED	SCALE
3319CA	Bain's Kloof	Topographic	1979	2	1:50000
3319	Worcester	Topographic	1980	3	1:250000
WEL_10	Wellington	Topographic	1985	1	1:500
3317	Cape Town	Admin	1982	1	1:500000

CONTOUR_INT	MAGIST_DIS	SOURCE
20 m	No	Surv. & L. Inf
50 m	No	Surv. & L. Inf
2 m	No	Wellingt.Mun.
100 m	Yes	Surv. & L. Inf

Fig. 4.8 The relation 'Map' in 1NF

This relation is then split to form the following two relations which are in 2NF:

MAP (REFERENCE, TITLE, TYPE, SCALE, CONTOUR\_INT,  
MAGIST\_DIS, SOURCE)

MAP\_EDITION (REFERENCE, DATE, ED)

This is demonstrated in Figure 4.9.

REFERENCE	TITLE	TYPE	SCALE
3319CA	Bain's Kloof	Topographic	1:50000
3319	Worcester	Topographic	1:250000
WEL_10	Wellington	Topographic	1:500
3317	Cape Town	Admin	1:500000

CONTOUR_INT	MAGIST_DIS	SOURCE
20 m	No	Surv. & L. Inf
50 m	No	Surv. & L. Inf
2 m	No	Wellingt.Mun.
100 m	Yes	Surv. & L. Inf

Fig. 4.9 The relation 'Map' in 2NF

A relation is in 3NF if and only if the non-key attributes (if any) are

- (a) mutually independent, and
- (b) fully dependent on the primary key.

The attribute MAGIST\_DIS of the relation MAP is dependent on TYPE which is a non-key attribute; this is known as transitive dependency and must be removed to create a relation of the "third normal form". This is

done by splitting the relation into two relations as shown in Figure 4.10.

The shorthand notations are as follows:

MAP\_DATA (REFERENCE, TITLE, TYPE, SCALE,  
CONTOUR\_INT, SOURCE)

MAP\_TYPE (TYPE, MAGIST\_DIS)

REFERENCE	TITLE	TYPE	SCALE	CONTOUR_INT	SOURCE
3319CA	Bain's Kloof	Topographic	1:50000	20 m	Surv.L.I.
3319	Worcester	Topographic	1:250000	50 m	Surv.L.I.
WEL_10	Wellington	Topographic	1:500	2 m	Wellingt.
3317	Cape Town	Admin	1:500000	100 m	Surv.L.I.

TYPE	MAGIST_DIS
Topographic	No
Admin	Yes

Fig. 4.10 The relations 'Map\_Data' and 'Map\_Type' are in 3NF

#### 4.5.2 View Integration

This is the process of merging the relations for each user view into a single set of relations in third normal form. Different user views may result in different relations which are very similar. These should be integrated so that no data is duplicated. In the previous example a relation MAP\_SOURCE (SOURCE, ADDRESS, TELEPHONE, PRICE) was formed. From the user

views for aerial photographs and co-ordinates, similar relations will be formed. This is combined in one relation **SOURCE** taking care that no data elements are lost in the process.

#### **4.5.3 Conceptual Data Model**

This model shows the relations and the relationships between them in graphic form. This can be done for all the user views to ensure that there is no duplication and to show how relations are used jointly by the different users. Figure 4.11 shows how the relations from *Map* can be joined to relations for aerial photographs.

In the relational data model, relationships between tuples (relations) are implicit, and are created by the use of "foreign keys". This is a non-key attribute in one relation that appears as the primary key (or part of the primary key) in another relation. The attribute **TYPE** is a foreign key in the relation *Map\_Data* (Figure 4.11) and implicitly creates the relationship with *Map\_Type*.

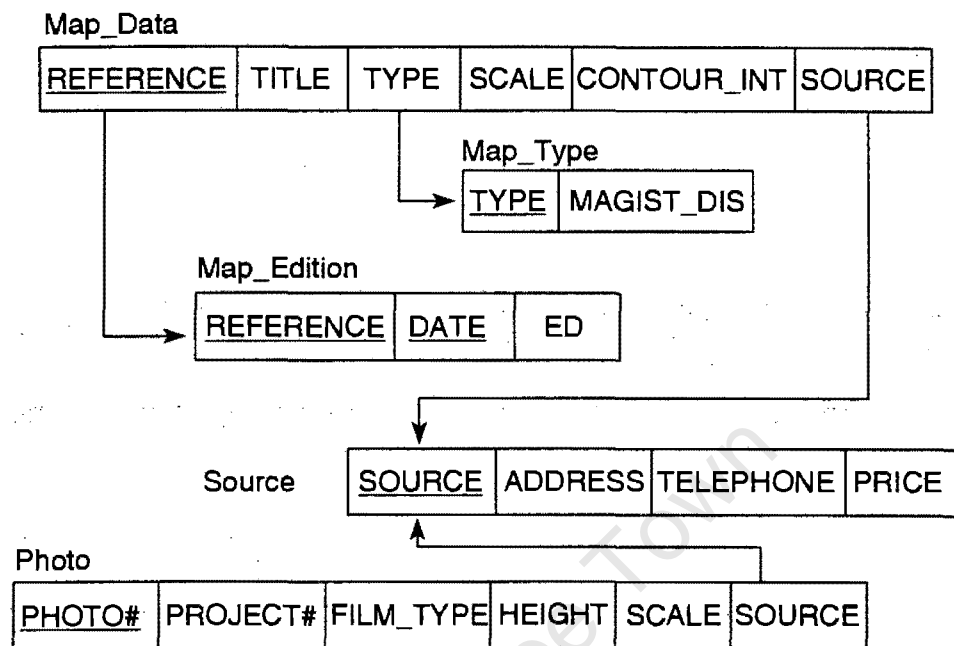


Fig. 4.11 A data model of the relations for maps and for aerial photographs

There is a one-to-many (1:m) relationship between *Map\_Data* and *Map\_Edition*, because every map may have a number of editions. This is established by including the key (REFERENCE) from the relation *Map\_Data* as a foreign key in the relation *Map\_Edition*. There is also a 1:m relationship between *Source* and *Map\_Data*, because an organisation may produce many different maps, therefore SOURCE is a foreign key in the relation *Map\_Data*.

To implement a many-to-many (m:n) relationship a new relation is created containing the primary keys of the related entity sets. This is illustrated in Figure 4.12 which shows a m:n relationship between *Erf* and *Map\_Data*.

For each erf there may exist a number of different maps, and on each map more than one erf may be shown. The primary keys of the participating relations become foreign keys in the relation *Erf\_Map*. The combination of both foreign keys form the primary key for the new relation, this is also known as a compound key.

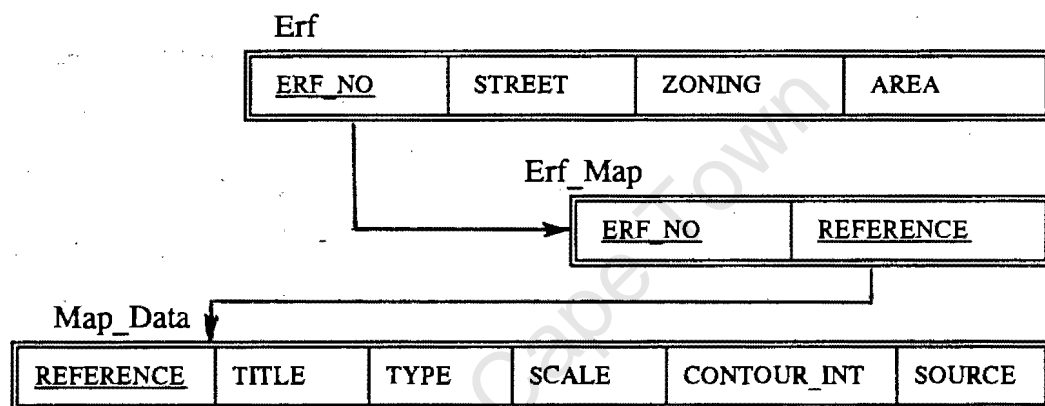


Fig. 4.12 A many-to-many relationship is implemented by creating a new relation 'Erf\_Map'.

The following tables were created for the alphanumeric database to store the metadata:

- *Erf* - stores basic data of all the erven and is the link between the geographical and alphanumeric databases.
- *Map* - will store information about all hard copy maps.
- *Map\_Date* - this table will make provision for different editions of the



same map.

- *Photo* - data about photographs and the film and camera used.
- *Orthophoto* - detail about orthophoto maps and a reference to the aerial photographs used to compile it.
- *Digital\_Data* - all data available in digital format, including maps and satellite imagery.
- *Co-ordinate* - co-ordinates of points that are marked.
- *Text* - any information in textual form e.g. reports and descriptions.
- *Data\_Source* - information about the supplier of the data.

The complete data model for the metadata base and the relations between the tables is given in Figure 4.13.



## **4.6 IMPLEMENTATION DESIGN**

### **4.6.1 Mapping to a Logical Model**

The most important step in implementation design is to map the conceptual data model into a DBMS-processible data model, which is called a "logical model" (also referred to as a schema). "Mapping" involves the transformation from one form or structure to another form or structure and the rules that govern this transformation. This model may have a hierarchical, network or relational structure. ReGIS, like most commercial GISs use a relational DBMS, and this makes the mapping process straightforward if the conceptual model is also in a relational format. Each relation in the data model becomes a relation in the database. Arrows in the conceptual data model are simply ignored, since they convey associations that are usually recorded as data within the relations.

### **4.6.2 Designing Applications**

A GIS is used to carry out analysis on the data and to give information; it must therefore be able to answer the following questions:

- What is at . . . ?
- Where is it . . . ?
- Which is the best way . . . ?
- What is the pattern . . . ?

- What if . . . ?

High-level languages or query languages such as SQL are used to ask these questions from the database and to generate summaries and reports.

#### 4.6.3 Physical Design

Physical design is the process of developing an efficient, implementable physical database structure. It is concerned with how the data is stored on physical devices rather than how it appears to the user. It involves the following:

- Design of stored records means the form of individual data records including pointers to other related records. It will make provision for the format of records, such as variable length, data compression to save storage space and record partitioning to improve data access time.
- Record clustering is the process of physically grouping records according to the dominant access paths, thereby minimizing access times.
- Design suitable access methods such as random or sequential search methods.

In a GIS this is normally accomplished by the database management system and the normal user does not have to concern himself with this. In a large installation a database administrator would be responsible for the physical design to improve the efficiency of the database.

## **CHAPTER 5**

# **BASE MAP AND INTEGRATION OF ATTRIBUTE DATA**

### **5.1 REQUIREMENTS**

A metadata base, as defined earlier, provides data about data. It is stored and accessed in an alphanumeric database, and although it may contain data about spatial objects or features, it does not have to be spatially related or accessible itself. The metadata base developed in the previous chapter is stored in a relational database and can be accessed and queried like any relational database.

A method whereby a graphical query is executed on a relational database is described by Fernández and Rusinkiewicz (1993) whereby the operator points at entities, relationships or attributes on the entity-relationship diagram. By selecting the required

entities and attributes the query is executed against the stored database. However, this does not make use of the spatial attributes, and it assumes that the user has some database knowledge of entities and E-R diagrams.

The requirements of the metadata base are to:

- provide the user with sufficient detail about spatial data available, to enable him to determine whether it would fulfil his requirements;
- supply information about where and how the data can be obtained;
- provide access to the metadata base in such a way that the general user, who may be computer illiterate, will be able to make use of it;
- contain data about all the spatially related data of the area, whether the actual data is spatial or not;
- encourage the data supplier to provide and maintain metadata of all his products.

Spatial data lends itself to visualisation and so it can also add a dimension to our cognition and hence to our means for searching on metadata. Because it is a spatial information system it would, therefore, make sense to make use of a spatial database and to also provide access via geographical position to the data. This would enable

the user not only to perform alphanumeric queries on the database, but also to do so via position and coverage. Some spatial data will be included in the database which will increase its value and make it more user friendly.

As a test area the town of Wellington was selected, because it is a medium-sized town having residential, business and industrial areas. A few blocks in the centre of the town was selected as an urban test area, and a block of farms just outside the town was used for a rural test area. The author has good relationships with staff in the City Engineer's department, which made it easy to obtain spatial data for the area. Another important reason was that Wellington Municipality was using ReGIS to develop their own GIS and digital data could be obtained in the correct format by the author for use in his ReGIS system.

## **5.2 BASE MAP**

### **5.2.1 Contents**

Once it was decided that spatial data must form part of the database, the format of this spatial data had to be determined. It would be used to form a base map or framework for the alphanumeric data. The purpose of a base map is to serve as backdrop for other spatial data. Features such as cadastral or administrative boundaries can not be seen on the earth, but by showing their relationships with visible features such as roads, railway lines and



rivers, their actual position can be visualised by the user.

Two factors had to be considered in setting up a spatial information system:

- the definition of the basic spatial unit; and
- the choice of referencing system.

The basic spatial unit is the smallest area for which data will be collected. For example, this could be a cadastral land parcel, a block in a township or a quarter-degree 1:50000 map sheet area. It should be the smallest identifiable spatial unit for which data can be collected consistently. In South Africa, with its excellent cadastral system, this should be an erf in municipal areas, and a farm or the smallest subdivision thereof, outside municipal areas.

In some undeveloped and remote areas there may be very large cadastral properties, e.g. national parks and tribal land. Especially the last may cover thousands of hectares and house many families, although it is only registered as one cadastral property. Will it still be viable to query such a large area with just one reference? Such areas would be regarded as undeveloped and there would exist much less spatial data for the area. The data that does exist would be small scale and therefore covering a large area, so that the cadastral property would still provide suitable access to the database.

A cadastral property can be represented on a map in one of the following ways:

- *Actual geometry* - the geometric figure is drawn to scale, e.g. an erf on a general plan.
- *Symbol* - where the property is represented by a symbol, e.g. where just the name of a farm is shown.
- *Containment* - where the scale of the map is too small to distinguish between individual properties although they fall within the area covered by the map.

The alphanumeric data in the metadata base will be related to the spatial data by means of referencing. Areal identification schemes differ. In addition to various co-ordinate systems, there are postal/street addresses, postcodes, and a variety of areal (often administrative) boundaries defined by various agencies. All of these, however, cover a large area. Erf numbers and farm numbers are unique reference systems which refer to the basic spatial unit and would be suitable as a reference system, because it is also the system whereby a cadastral property is identified.

In Paragraph 3.3 it was shown that the biggest demand for spatial data is for topographic, political and cadastral data. The function of topographic data is to locate a position in space and to use this information for the planning of development and infrastructure. Topographic data is also important so that a user can orientate himself on the map and locate the area of interest. Political data is provincial, magisterial and municipal boundaries, and this is used mainly for administrative purposes. Cadastral boundaries are the demarcation of properties and is the basic land unit in South Africa. This often indicates the limit of any spatial investigation, planning or development that takes place.

When setting up a base map it would make sense to include some of this data that are frequently required by the users as the base data, because it would add to the value of the database, since this is not only metadata, but real data which may be used directly.

It was decided not to show relief data such as contours on the base map because this would make the map appear cluttered and would not improve the readability of the map. In the urban area the block and erf cadastral data would be sufficient for the user to identify his area of interest, while in the rural area the roads, railway lines, rivers and farm boundaries would be sufficient for this purpose.

### 5.2.2 Acquisition

Another important factor that had to be considered was the acquisition of geographical data for the base map. Two options could be considered: digitizing from maps or acquiring existing digital data. An important factor throughout this research has been to endeavour to reduce the amount of duplication in data and data capture. If suitable spatial data for the base map could therefore be obtained, this should be used.

Because the use and the availability of spatial data, as well as the type of features, are different for urban and rural areas, these have been considered separately. For urban areas the most important data would be cadastral boundaries, while topographic features such as contours, rivers and vegetation would not be of any significance for the base map. This cadastral data for the urban test area was obtained in structured GIS format from the Wellington Municipality. It contained the geographical data and attribute data for the erven such as erf number, area, zoning, street name and number. The erf number was the key field and the link between the spatial and alphanumeric databases.

For the rural test area near Wellington the standard digital 1:50000 database from the Directorate of Surveys and Land Information was used. This consisted of the communication features such as roads, railways, rivers and built-up areas, but very limited cadastral information. It showed only the

original farm boundaries, which have since been subject to a considerable amount of subdivision and consolidation. The cadastral data provided, therefore, being out-dated, was insufficient for the requirements. The office of the Surveyor-General is responsible for capturing cadastral data, but is still engaged in this process. The rural area around Wellington was only available as a CAD drawing showing cadastral boundaries originally digitized from noting sheets. This had to be converted into GIS format which involved classifying it by adding intelligence and topology.

It was decided not to show any contours, because this would be confusing rather than helpful to the general user. Too much detail would not assist the recognition of areas.

## **5.3 INFORMATION INTEGRATION**

### **5.3.1 Definition**

A GIS brings information together, it unifies and integrates that information. It makes available information to which no one had access before, and places old information in a new context. It often brings together information which either was not or could not be brought together previously.

Information integration is the synthesis of geographical information in a

computer system which depends for its effectiveness on information linkage (i.e. of spatial and attribute data) within a coherent data model. This involves bringing together diverse information from a variety of sources, which means information interchange.

### **5.3.2 Integration of Spatial and Attribute Data**

In commercial information systems, non-spatial data are typically integrated by storing them non-redundantly in a single database engine. In computer mapping systems, spatial data of various kinds are frequently integrated within a single graphical database, in either vector or raster format. Within each of these systems, the relationship between items of consistently recorded data can be used to answer a variety of questions. However, the range of questions that can be asked must be of a spatial or non-spatial nature.

Within GIS, information integration goes one step further, involving the linkage of non-spatial or attribute data to spatial information describing real-world features. By performing operations across the two sets of information in tandem, a far richer set of questions may be asked, and a far broader range of problems can be solved than in those systems that handle just attribute or spatial data alone.

In most vector systems the link between spatial and attribute information is created by the geo-relational model. This is done by arranging for each spatial feature's unique identifier (or ID) to be recorded in a key field of the appropriate database table(s) that stores its attribute information. A database join can work both ways: attribute information can be found by selecting spatial features, and spatial features can be found by querying attribute information. Furthermore, all attribute information within the model is associated with one or more spatial features.

Most vector-oriented GISs usually adopt a dual or hybrid data storage strategy to implement the model, with spatial data held separately from the attribute data, and the system maintaining links between the software modules that handle each information type. Locational data is typically represented using a topological (or arc-node) spatial data model implemented by specialised proprietary data structures, and thematic data is usually stored in the tables of a standard relational database. The use of a standard relational database for managing attributes provides flexible access to this data and the possibility of integrating attribute data with other applications. In general, significant benefits can be gained by storing all of an organisation's data in a single DBMS. By doing this it is possible to easily link different kinds of data from different departments, especially when using the relational model. From an initial situation where each GIS supported only one DBMS, a clear trend is now being established to provide access to multiple DBMSs such as Oracle, Sybase and Informix.

## **5.4 UPDATING THE GEOGRAPHICAL DATA**

The main function of the geographical data in a metadata base system is to serve as base map, and not to act as a source of information. This must be kept in mind when setting up a policy of updating the geographical data. The cadastral and topographic data have been obtained from different sources, and will be discussed separately.

### **5.4.1 Cadastral Data**

The alphanumeric database is linked to the geographical base map with keys which are connected to erven or cadastral land parcels. The attribute data in the metadata base is unique for every land parcel. It was shown in Paragraph 5.2 that the user will access the data via a cadastral land parcel. Cadastral data will change in the following two cases: if a property is subdivided, or consolidated with one or more adjacent properties.

In the case of subdivision, all the attribute data that was related to the original property, will also apply to all its subdivisions. If the data is not updated, the original data will still be correct. When an erf is subdivided, new erf numbers are created for the subdivisions, while the remainder retains the original number, for which the data is still correct. For the new subdivided erven the data in the database on e.g. maps, photographs and digital data must be added. This will be the same as for the original erf.



When properties are consolidated, a new erf number is allocated to the property. This must be added to the database with attribute data which were common to all the portions that make up the new property. Attribute data that applied only to some of the component properties can not be included because it does not apply to the consolidated property as a whole. This can be included with a note that it applies only to a portion of the property.

#### **5.4.2 Topographic Data**

The topographic data consists of roads, railways and powerlines which are mainly used for orientation purposes so that the user can locate the area of interest. These data are stable and do not change frequently and will only have to be updated when major new roads or railways are built. A note will inform the user that this data is only serving as a base map, and that the updated maps must be obtained from the supplier whose data will appear in the metadata.

## **CHAPTER 6**

# **RETRIEVAL OF METADATA**

### **6.1 INTRODUCTION**

The purpose of a metadata information system is to supply the users with information about spatial data. It must manage and display data about data sets and non-digital data so that users can find the information they need to help them solve their problems. Such a system should assist a user to build up a more accurate mental model of the available data which are likely to be useful for their purposes.

The particular aspect of spatial meta-information systems (SMIS) that sets them apart from other meta-information systems is that information may be held and retrieved on the basis of a spatial reference. Traditional query languages are restricted to the

treatment of properties of non-spatial data. A SMIS, on the other hand, must provide a way for the user to enquire about spatial data in a spatial way. Thus a SMIS differs from conventional information systems in the specific interaction between spatial and non-spatial objects.

The metadata is stored as part of a GIS to make it more flexible. This, however, means that the data is not easily accessible for the general user who is not familiar with the particular GIS. A system must therefore be developed to enable both experienced and general users to retrieve the data. They must be able to retrieve and present the data in a suitable format without learning the GIS functionality or SQL commands to access the attribute data in the alphanumeric database.

## 6.2 REQUIREMENTS

The primary aim of the system is to make it easy for users to find out what data exist that may help them find the answer to some problem, and where any relevant data are held. If the system is to be really useful, users must believe that the information provided in answer to a query is both complete and up-to-date. In addition to being informed about the existence of data, users may require information about the quality, cost, accessibility and format of the data. The system should be able to assist the user even if data from a number of different data sets or data types would be required fully to answer a query. Furthermore, it is expected that people from many different backgrounds may use the system so no assumptions can be made about the interests

or expertise of the users when they make enquiries.

### 6.2.1 Diversity of Data

The data may vary widely in a number of different ways, such as:

- *Type of data*, e.g. map, picture (raster or vector, monochrome or colour), text, numeric or sound.
- *Storage medium*, e.g. magnetic disc, CD-ROM, video cassette, paper or microfiche.
- *Storage location*, e.g. state, local authority or private organisation.
- *Storage organisation*, e.g. structured (relational database, indexed file) or unstructured.
- *Volatility*, e.g. static, infrequent changes or regular changes (annual or monthly).
- *Availability*, e.g. at any time, only at specified times or only with prior notice.
- *Cost*, e.g. free, single payment or charged by amount of information

obtained.

- *Accessibility*, e.g. available to anyone, available only to people on a restricted list, or available only under specified conditions (such as to an employee).
- *Degree of processing* (i.e. how much has the data been processed), especially data aggregation.
- *Quality* for a given purpose: e.g. the quality of earth observation data could be assessed relative to geometric accuracy, the presence of gross errors or deficiencies, and instrument measurement accuracy.

The basic data items contained within one data set can be very different from those contained in another. Examples of individual data items might be:

- 1:50000 topographic map reference 3319CA;
- the number of people living in a counting area during the 1990 census;
- the grazing capacity of an agricultural area.

## **6.3 DATA RETRIEVAL**

Medyckyj-Scott (1991) described two spatial meta-information systems: those built on Database Management Systems (DBMS) and those built on Information Retrieval Systems (IRS) technology. DBMS is primarily concerned about storage efficiency, continuous amendment of data by many different users, and security. They handle numeric and text data in tabular form which are highly structured and tend to be of a fixed format and length. These data are usually factual and retrieved in response to a well-defined query in order to provide a concrete answer to a specific problem.

Conventional IRSs are defined as systems which provide facilities for storage of and access to information about documents. Unlike DBMS, retrieval is not based upon readily identifiable keys such as personal names or code numbers, but primarily on textual information.

A spatial meta-information system based on a GIS makes use of a system where the spatial and alphanumeric databases are combined, and therefore has direct access to the technology of both worlds.

### **6.3.1 Dialogue Styles**

The following are different forms of query dialogue styles. These can be considered to vary on the degree to which the interaction is system- or user-driven.

- *Line by line.* This dialogue consists of the display of prompts and information by the system with the entry of commands by the user. Searches are performed by a query language such as SQL, and the user must therefore have a working knowledge of this query language.
- *Queries constructed by the use of form-filling.* A formatted screen is used to collect the commands, the database parameters and their relationships. The user constructs a query statement by selectively filling in fields with key words or by entering parameter values to restrict information access in accordance with their requirements. ReGIS makes provision for this type of query. The disadvantage of this system is that the user must know the format of the database.
- *Query by example* was developed to operate on a relational model of data. It is a special case of form-filling, with users filling in fields in tables to specify record type, the fields to be searched, the retrieval criteria, what is to be extracted and how.
- *Hierarchical menu selection.* With this method, users make a choice from a menu of options, possibly keywords, offered by the system. The menus may be structured hierarchically with the choice of an option resulting in the display of a new menu. This is one of the easiest ways for users to formulate query statements, but will be confusing if the keywords are not very carefully chosen. It also

becomes tedious for an experienced user.

- *Querying using graphical query languages.* With these the user formulates requests mainly in terms of diagrams but some typed input is usually required. In spatial meta-information systems, graphic querying can be used to obtain information from the user as to the spatial area for which information is required.

While each of the above forms of querying has its advantages and disadvantages, the type of interaction dialogue provided for a SMIS must be based on consideration of the principal types of users and the needs that the system is intended to fulfil.

### 6.3.2 Output

In order to aid comprehension, the result of a query should be output in a manner that is desirable to the user. The designer should be concerned about: how much information is presented, in what form and what order; how the user should control the format of the output, and whether there will be a need for hard copy output.

We can distinguish between users scanning information and reading it in detail. Typically, the results of a search will be shown in the form of a summary list. During the initial evaluation of the results, the user compares



the original request with the information that is being returned by the system. A large number of unsuccessful queries are related to factors such as the content of the databases and the way in which the information is structured and stored. Apart from the system not having the information the user requires, poor results normally result from queries containing too broad or too narrow search phrases, incorrectly used operators, incorrect use of logic or the misspelling of terms.

Although special care should be taken to ensure the correctness and currency of the data, this is never completely possible due to a number of factors:

- There is always a time lag between the creation or amendment of a feature on the earth and its capture on a central database.
- Not all suppliers will supply metadata due to security or other reasons.
- Some suppliers may not be aware of the existence of such a catalogue or its advantages.
- Notwithstanding all possible checks that may be performed on the data, mistakes do slip in during data capture or transfer.

A disclaimer must therefore be given by the system stating that neither the agency providing the service nor the supplier of the data can be held responsible for any errors or omissions in the data.

## 6.4 DESIGN CRITERIA

The following principles should be considered when designing a spatial meta-information system:

- *Look to a user-centred design.* This means that the user requirements must be considered during the design process. What their information needs are and what the system will have to do to meet these needs.
- *Consider the characteristics of the users.* The users of a SMIS - both data suppliers and data users - will cover a wide range of skills and experience; many will be computer experts but many will have little experience. Most users will be intermittent users. The user interface must therefore be easy to operate again, even after a considerable period of non-use by any one individual.
- *Reduce the likelihood of user misconceptions.* Examples are that the terms used by the user do not have the same meaning to the system, or that the query is too broad or too narrow.

- *Provide flexibility for handling the spatial component of a query.* This can be done by indicating the area of interest or by defining the geographical or grid co-ordinates.

#### 6.4.1 Menu Design

For general users a menu provides the easiest and most user friendly form of access. It also creates the possibility that the experienced user could move quickly through the menus. According to Medyckyj-Scott (1991) an important consideration is to create a user interface that has the qualities of simplicity, a standard language (in lieu of jargon), and avoidance of deep menus (that is, the necessity to make many choices to get to one's desired result).

An important choice the user would have to make is whether he wants to query the database in a visual manner, i.e. by pointing at the area of interest, or by locating it by reference, i.e. erf or farm number. In either case access to the metadata is provided via a particular property. The other important decision one has to make is the type of data required, e.g. maps, photographs, digital data, orthophoto maps or text.

Persons familiar with the GIS would be able to have direct access to the database and could make use of the GIS functionality to gain faster access to the metadata, and even to perform analysis on the data.

## 6.5 ACQUIRING METADATA

This problem is not one that has been considered very much until now, because as stated by Newman (1991) systems need to exist before the problem of filling them becomes apparent. Experience from digital data has shown that someone must be explicitly paid to carry out the task, or people must recognise that it is in their interest to supply the information. Neither of these criteria necessarily apply when the required information is scattered in many locations and each individual controls or has access to only a small part of the total. From an individual contributor's viewpoint, it can be seen that unless other people can be persuaded to supply information it will not be worth the labour of generating one's own contribution because the information in the system will be incomplete. Furthermore, even if users supply information there is still the question of getting it into the system and of checking it.

Newman (1991) suggested that the only practical method is:

- to persuade users to enter their own metadata;
- to design the system so that it makes as many automatic checks on the validity of data as possible; and
- to accept that there will always be some erroneous data in the system.

In fact, the only way to get started might be for some organisation to provide and pay for the entry of the metadata that they need. If the information that was entered was also of interest to other users then they could, perhaps, be persuaded to contribute to the costs of data entry by paying a fee for the data that was returned in answer to their queries. The overall effect would then, hopefully, be a snowball, the more information that is available the greater the probability that people would use the service. Similarly, the more people that use the service the greater the likelihood that depositors would see a value in supplying their own meta-information.

## **CHAPTER 7**

### **CONCLUSIONS**

#### **7.1 SPATIAL DATA SUPPLY AND DEMAND**

It has been shown that information is the most important commodity for managers and decision-makers. Without the correct information, informed decisions can not be made. Especially for land management, suitable spatial information must be available and accessible to decision-makers and other users.

Information has some unique features:

- it does not get used up;
- it can lose value even if it is never used. Currency and completeness is

therefore very important;

- the same information can be used by many users at the same time and for different purposes; and
- it is inexpensive to reproduce, even though the original collection of the data may have been expensive;

It was found in the survey that a large number of spatial data types were produced by the suppliers and that there were no major deficiencies in the supply. A number of users, however, were not aware of all the data types available, for example maps of neighbouring countries, archaeological maps, land use maps for urban and rural areas, soil types and recent aerial photographs. Most of these are available, but the users are not aware of this, or where to obtain them.

The problem of data supply would therefore not be solved by producing more data, because this would just increase the duplication of data. Rather it should be ensured that the available data reaches the user. A method of bringing the data suppliers and users together should be found.

## 7.2 DATA DISTRIBUTION

Spatial data suppliers and users can be grouped into four categories:

- *Spatial data managers:* they are the major suppliers of spatial data and include air survey companies, cartographers and surveyors.
- *Planners and construction:* this group consist of engineers, town planners and architects and are the major users of spatial data. They normally use the basic data and enhance it by adding information to it.
- *Land use and development:* these include ecologists, geologists, farmers and property developers and are mostly users of spatial data. They are not aware of the variety of spatial data available or where to obtain it, but they do have access to the essential data they require. They are also producers of some unique data such as the distribution of certain bird species or the potential and soil type of agricultural land.
- *Services and information:* this is provided by libraries, educational institutions, TELKOM and the CSIR. Many individuals in these organisations have no formal training or experience in the use of spatial data, but they have to make use of spatial data to perform their duties. Their knowledge and use of spatial data is restricted to common types which they regularly use.



To expand the market for spatial data, the suppliers must specifically look at the last two categories and at ways to reach them.

### **7.3 METADATA BASE**

Three possible solutions were considered to improve the distribution of spatial data and to reduce the duplication of data:

- a comprehensive GIS with a full database;
- a national spatial data infrastructure; and
- a metadata base.

A comprehensive GIS is ideal when an organisation has the hardware and expertise to operate and make use of the analysis functionality. In an environment where many users do not have access to the hardware, this would not provide access to spatial data for them. A national spatial data infrastructure has the same disadvantages and would only operate effectively amongst users which have similar hardware and database installations.

On the other hand, a metadata base would also provide access to users who do not have sophisticated computers available. It will provide them with the necessary

information to obtain the spatial data they require. A network can be set up for users with the necessary equipment to access the database on-line. This would also enable data suppliers with an easy way to update the database with their information.

## 7.4 DATA SHARING

The success of data sharing depends on the completeness of the metadata base. It is therefore essential that a method is developed to ensure the participation of all data suppliers in supplying and maintaining data.

An example can be borrowed from the property market. An agent would set up such a metadata base and act as an agency. He would collect the data from the suppliers and maintain the database. He would also be responsible for advertising this information system amongst possible users. Once the user has obtained metadata about suitable spatial data that would fulfil his requirements, he would buy it directly from the supplier and pay a commission to the agent.

The commission would be the incentive to the agent to provide a good information system and advertise it as widely as possible so that all possible users could be informed about its existence. The supplier would be motivated by the profit on his sales, to supply the metadata base with the latest information about his data (and thus maintaining it). Once the system is in place it should have a snowball effect, and the success would hopefully convince other suppliers and users to take part.

The system should be provided on a regional basis because in this manner the agent would be better informed about the suppliers and users of spatial data and their needs. The local authority, which is also a major supplier and user of spatial data, could serve as a venue for the information system.

The properly maintained metadata base, with its access to information being constantly and regularly updated, will enable key decision makers to make the best informed decisions over the broadest possible spectrum.

## BIBLIOGRAPHY

Adlam K A McL, Clayton A R and Kelk B, 1988. A 'demonstrator' for the National Geosciences Data Index. *International Journal of Geographical Information Systems*, 2(2), p.161-170.

Akervall L, Degerstedt K and Rystedt B, 1991. Spatial metadata systems at the National Land Survey of Sweden. In *Metadata in the Geosciences*, Medyckyj-Scott D, Newman I, Ruggles C and Walker D (ed.). Loughborough: Group D Publications.

Al-Taha K and Frank A, 1991. Temporal GIS Keeps Data Current. In *1991-92 International GIS Sourcebook*, p.384-388.

Anderson K E and Callahan G M, 1990. The modernization program of the U.S. Geological Survey's national mapping division. *Cartography and Geographic Information Systems*, 17(3), p.243-248.

Aronoff S, 1989. *Geographic Information Systems: A Management Perspective*. Ottawa: WDL Publications.

Batty P, 1992. Exploiting relational database technology in a GIS. *Computers & Geosciences*, 18(4), p.453-462.

Bauer K W, 1994. Creation of Land Information Systems for Milwaukee County, Wisconsin. *Journal of Surveying Engineering*, 120(4), p.156-165.

Beck F J and Olsen R W, 1986. Quality control and standards for a national digital cartographic data base. *Auto Carto London*, Vol. 1, p.372-380.

Bossler J D, Finnie T C, Petchenik B B and Usselman T M, 1990. Spatial data needs: The future of the national mapping program. *Cartography and Geographic Information Systems*, 17(3), p.237-242.

Buehler K, 1994. OGIS Augments Data Transfer. *GIS World*, 7(10), p.60-63.

Burnhill P, 1991. Metadata and cataloguing standards: one eye on the spatial. In *Metadata in the Geosciences*, Medyckyj-Scott D, Newman I, Ruggles C and Walker D (ed.). Loughborough: Group D Publications.

Burrough P A and Frank A U, 1995. Concepts and paradigms in spatial information: are current geographical information systems truly generic? *International Journal of Geographical Information Systems*. 9(2), p.101-116.

Cassettari S, 1993. *Introduction to Integrated Geo-Information Management*. London:

Chapman & Hall.

Cassettari S, 1994. One-stop shopping for spatial data? *Mapping Awareness*, 8(3), p.58.

Cassettari S, Fagg A and Visvalingam M, 1992. Cartography and Geographical Information Systems. *The Cartographic Journal*, 29(1), p.51-54.

Charlwood G, Moon G and Tulip J, 1987. Developing a DBMS for Geographic Information: A Review. *Proceedings Auto Carto 8*, p.302-315.

Chorley R, 1988. Some reflections on the handling of geographical information. *International Journal of Geographical Information Systems*, 2(1), p.3-9.

Clarke D G, 1989. The national land information system and the land surveyor in South Africa. *Proc. Ninth Conference of Southern African Surveyors*, 10 pages.

Clarke D G, 1995. Personal communication.

Clarke K C, 1990. *Analytical and Computer Cartography*. Englewood Cliffs: Prentice-Hall.

Coleman D J and McLaughlin J D, 1994. Building a Global Spatial Data Infrastructure: Usage Paradigms and Market Influences. *Geomatica*, 48(3), p.225-236.

Cooper A K and Clarke D G, 1991. The South African Standard for the Exchange of Digital

Geo-referenced Information. In *Spatial Database Transfer Standards: Current international status*, Moellering H (ed.). Elsevier.

Coppock J T and Rhind D W, 1991. The history of GIS. In *Geographical Information Systems: principles and applications*, Maguire D J, Goodchild M F and Rhind D W (eds.). London: Longman.

Corbley K, 1995. Citizens enjoy GIS access. *GIS World*, 8(3), p.66-68.

Cromley R G, 1992. *Digital Cartography*. Englewood Cliffs: Prentice-Hall.

Cuthbertson M, 1993. A Database for Environmental Research Programmes. *Journal of Environmental Management*, 37, p.291-300.

Dale P F and McLaughlin J D, 1988. *Land Information Management*. New York: Oxford University Press.

Dangermond J, 1989. Starting GIS technology: Three important processes to remember. *Public Works*, 120, p.42-44.

Dangermond J, 1992, *What is a Geographic Information System (GIS)?*, In Johnson A I, Pettersson C B and Fulton J L (Eds.), *Geographic Information System (GIS) and Mapping - Practices and Standards*, American Society for Testing and Materials, p.11-17.

Date C J, 1990. *An Introduction to Database Systems*. Vol. 1, 5th ed., Reading: Addison-Wesley.

Densham P J, 1991. Spatial decision support systems. In *Geographical Information Systems: principles and applications*, Maguire D J, Goodchild M F and Rhind D W (eds.). London: Longman.

Donahue J G, 1990a. AM/FM - GIS/LIS in the 21st. century. *Proc. ACSM-ASPRS Annual Convention*, Vol. 3, p.58-65.

Donahue J G, 1990b. Land base accuracy: Is it worth the cost. In *Geographic and Land Information Systems for practicing surveyors*, Onsrud H J and Cook D W (Eds.). Bethesda: ACSM.

Durgin P M, 1993. Measurement-Based Databases: One Approach to the Integration of Survey and GIS Cadastral Data. *Surveying and Land Information Systems*, 53(1), p.41-47.

Dymon U J, 1989. Do we really know our map users? *Cartographica*, 26(3 & 4), p.49-58.

Elmasri R and Navathe S B, 1989. *Fundamentals of Database Systems*. Redwood City, California: Benjamin/Cummings Publishing.

Eyre L A, 1989. JAMGIS, the first Jamaican Government comprehensive multi-data geographical information system. *International Journal of Geographical Information Systems*,



3(4), p.363-371.

Fairbairn D J, 1994. The Frontiers of Cartography: Mapping a changing discipline. *Photogrammetric Record*, 14(84), p.903-915.

Fernández R N and Rusinkiewicz M, 1993. A conceptual design of a soil database for a geographical information system. *International Journal of Geographical Information Systems*, 7(6), p.525-539.

Fisher P F, 1991. Spatial data sources and data problems. In *Geographical Information Systems: principles and applications*, Maguire D J, Goodchild M F and Rhind D W (eds.). London: Longman.

Flowerdew R, 1991. Spatial data integration. In *Geographical Information Systems: principles and applications*, Maguire D J, Goodchild M F and Rhind D W (eds.). London: Longman.

Frank U A, 1988. Requirements for a Database Management System for a GIS. *Photogrammetric Engineering & Remote Sensing*, 54(11), p.1557-1564.

Goodchild M F, 1990. Geographic Information Systems and cartography. *Cartography*, 19(1), p.1-13.

Grant D M, 1993. Provision and Management of Spatial Data in New South Wales. *Proc. LIM & GIS conference*, Sydney, p.95-109.

Green D R, 1993. Map output from geographic information and digital image processing systems: a cartographic problem. *The Cartographic Journal*, 30(2), p.91-96.

Greve C W, Kelmelis J A, Fegeas R, Guptill S C and Mouat N, 1993. Investigating U.S. Geological Survey needs for the Management of Temporal GIS Data. *Photogrammetric Engineering & Remote Sensing*, 59(10), p.1503-7.

Groot R, 1992. Management aspects of the introduction of spatial information technology in mapping organizations. *ITC Journal*, 1992-4, p.353-358.

Guptill S C, 1986. A new design for the U.S. Geological Survey's national digital cartographic data base. *Auto Carto London*, Vol. 2, p.10-18.

Guptill S C, 1992. A national geographic data system: a concept. *Proc. GIS/LIS '92*, Vol. 1, p.300-302.

Harlow C, 1989. The future of geographic information systems: 18 important trends. *Proc. ASPRS/ACSM Annual Convention*, Vol. 5, p.16-25.

Haywood P, 1986. The Differing Requirements of Producers and Users and the Design of Topographic/Cartographic Databases. *The Cartographic Journal*, 23(2), p.131-134.

Healey R G, 1991. Database Management Systems. In *Geographical Information Systems: principles and applications*, Maguire D J, Goodchild M F and Rhind D W (eds.). London:

Longman.

Healey R G and Morris B A, 1987. CARTO-NET: a relational database approach to automated map cataloguing. *Cartographic Journal*, 24, p.15-18.

Hiland M W, McBride R, Davis D, Braud D, Streiffer H, Jones F, Lewis A and Williams S, 1991. Louisiana Coastal GIS Network: Graphical User Interface for access to spatial data. *Proc. GIS/LIS '91, American Society for Photogrammetry and Remote Sensing*, Vol. 2, p.845-856.

Hiland M W, Wayne L and Streiffer H, 1992. Louisiana Coastal GIS Network: relational database design for a spatially indexed cataloging system. *Proc. GIS/LIS '92, American Society for Photogrammetry and Remote Sensing*, Vol. 1, p.322-338.

Holmes D O, 1990. Computers and Geographic information access. *Meridian*, 4, p.37-49.

Howe D R, 1983. *Data Analysis for Data Base Design*. London: Edward Arnold.

Hutt G, 1992. Maps retroconversion at the British Library. *Mapping Awareness & GIS in Europe*, 6(10), p.8-9.

Huxhold W E, 1991. *An Introduction to Urban Geographic Information Systems*. New York: Oxford University Press.

Kasturi R, Fernandez R, Amlani M and Feng W, 1989. Map Data Processing in Geographic Information Systems. *Computer*, 22(12), p.10-21.

Keates J S, 1992. *Understanding Maps*. London: Longman.

Kuggeleijn R, 1995. Managing data about data. *GIS Europe*, 4(2), p.32-33.

Lai P and Gillies C F, 1991. The impact of geographical information systems on the role of spatial data libraries. *International Journal of Geographical Information Systems*, 5(2), p.241-251.

Land Survey Act, Act 9 of 1927.

Langran G, 1992. *Time in Geographic Information Systems*. London: Taylor & Francis.

Lanter D P, 1991. Design of a lineage-based meta-data base for GIS. *Cartography and Geographic Information Systems*, 18(4), p.255-261.

Laurini R and Thompson D, 1992. *Fundamentals of Spatial Information Systems*, London: Academic Press.

Laursen V W, 1990. Utilities - a part of a Land Information System. *Proceedings of the XIX FIG Congress*, p.158-169.

Leatherdale J D, 1994. Map making for profit and pleasure. *Photogrammetric Record*, 14(83), p.759-767.

Lester K J, 1993. The cadastral and topographic components of the South African National Land Information System. *Proceedings of the Tenth Conference of Southern African Surveyors*. 11 pages.

Lester K J, 1994. The South African National Land Information System. *Proceedings of the 20th FIG Congress*, Vol. 3, 8 pages.

Lester K J, 1995. A Dynamic approach towards Integrated Land Information: Corporate strategy for a new South Africa. *Proceedings of GIS'95 Government Information Systems, Singapore*. 6 pages.

Lillywhite J, 1991. Identifying available spatial metadata: the problem. In *Metadata in the Geosciences*, Medyckyj-Scott D, Newman I, Ruggles C and Walker D (ed.). Loughborough: Group D Publications.

Loudon T V, 1992. GIS and the geological survey map. *Mapping Awareness & GIS in Europe*, 6(2), p.14-18.

Maffini G, 1990. The role of public domain databases in the growth and development of GIS. *Mapping Awareness*, 4(1), p.49-54.

Mahoney R P, 1986. Digital mapping - an information centre. *Proc. Auto Carto London*, Vol. 2, p.190-199.

Margeot M Y G H and Forbes A M, 1991. Geo-Map: A computer-based map cataloguing system. *Proceedings EDIS/SAGIS '91 Conference*.

Mawdsley S, 1992. The design and implementation of a computerized map library cataloguing system. *The Bulletin of the Society of Cartographers*, 26(2), p.13-16.

McArthur J A and McGrath G, 1991. Some current issues in cartographic distribution and future prospects. *The Cartographic Journal*, 28(2), p.200-207.

McBride R A, Davis D, Jones F, Byrnes M, Braud D, Hiland M, Lewis A and Streiffer H, 1991. Louisiana Coastal Geographic Information System Network: Access to spatial data. *Meridian*, 6, p.29-43.

McBride R A, Hiland M, Jones F, Braud D, Wayne L, Streiffer H, Carpenter M, Lewis A, Mogli S, Arnold J, Lingineni S and Williams S, 1993. Louisiana coastal Geographic Information System network: More than a spatially indexed cataloging system. *Proc. ACSM / ASPRS Annual Convention*, Vol. 3, p.238-255.

McFadden F R and Hoffer J A, 1988. *Data Base Management*. 2nd ed., Menlo Park, California: Benjamin/Cummings Publishing.

McLaughlin J and Anderson R, 1994. Access to Land-Related Information: Report of a working group. *Proceedings of the 20th FIG Congress*, Vol. 3, 10 pages.

McLaughlin J D and Nichols S E, 1987. Parcel-Based Land Information Systems. *Surveying and Mapping*, 47(1), p.11-29.

McLaughlin J and Nichols S, 1994. Developing a National Spatial Data Infrastructure. *Journal of Surveying Engineering*, 120(2), p.62-76.

Medyckyj-Scott D, 1991. User-oriented enquiry facilities. In *Metadata in the Geosciences*, Medyckyj-Scott D, Newman I, Ruggles C and Walker D (ed.). Loughborough: Group D Publications.

Monmonier M, 1993. *Mapping it Out*. Chicago: University of Chicago Press.

Morehouse S, 1992. The Arc/Info Geographic Information System. *Computers & Geosciences*, 18(4), p.435-441.

Morris B, 1986. CARTO-NET. British Library research project on map library automation at the University of Edinburgh Department of Geography. *The Bulletin of the Society of University Cartographers*, 20(1), p.31-32.

Morris B, 1990. Automated Search and Retrieval. In *Information Sources in Cartography*, Perkins C and Parry R B (Eds.). New York: Balkan.

Morrison A, 1979. The frequency of map use for planning journeys by road. *Society of University Cartographers Bulletin*, 13(1), p.35-41.

Newman I, 1991. Data Dictionaries, Information Resource Dictionary Systems and Metadatabases. In *Metadata in the Geosciences*, Medyckyj-Scott D, Newman I, Ruggles C and Walker D (ed.). Loughborough: Group D Publications.

Nichols S and McLaughlin J, 1993. Building a national spatial data infrastructure. *Proceedings of the Tenth Conference of Southern African Surveyors*. 7 pages.

Oberg R E, 1988. The visual Cartographic enquiry system for geodetic control data. *Proc. GIS/LIS '88*, p.484-489.

Oberg R E, 1989. The development of the visual cartographic enquiry system. *Proc. GIS/LIS '89*, p.301-304.

Ottens H F L and Zandee R H, 1991. The NexpRI Information Bank: a meta-information system on GIS activities in the Netherlands. In *Metadata in the Geosciences*, Medyckyj-Scott D, Newman I, Ruggles C and Walker D (ed.). Loughborough: Group D Publications.

Parr D M, 1990. A comprehensive data model for Local Government. *Proceedings of the URISA annual conference*, Vol. 3, p.99-110.

Pearman H, 1993. Designing a land and property gazetteer. *Mapping Awareness & GIS in*



Europe, 7(5), p.19-21.

Pearson A and Sprunt B, 1986. A microcomputer based map retrieval system. *The Bulletin of the Society of University Cartographers*, 20(2), p.100-102.

Peled A and Adler R, 1993. A common database for digital mapping and GIS. *International Journal of Geographical Information Systems*, 7(5), p.425-434.

Perkins C, 1992. Metaphysical mayhem? Retrieving and describing maps and spatial data in the map library and drawing office. *The Bulletin of the Society of Cartographers*, 26(2), p.21-24.

Perkins C, 1994. Quality in map librarianship and documentation in the GIS age. *The Cartographic Journal*, 31(2), p.93-99.

Perkins C et al, 1988. "All about maps": Establishing a CARTO-NET based map database. *The Bulletin of the Society of University Cartographers*, 21(2), p.88-89.

Porter J R and Lores P W, 1990. The South Australian Land Information System - practical applications for all. *Proceedings of the XIX FIG Congress*, p.213-225.

Poynter A, 1993. A digital map production system for a map publisher. *Mapping Awareness*, 7(9), p.26-27.

Price S, 1995. Moves towards national land information systems in South East Asia. *Proceedings of the FIG Symposium on Engineering Surveying*, Cape Town, 13 pages.

Raper J F and Maguire D J, 1992. Design models and functionality in GIS. *Computers and Geosciences*, 18(4), p.387-394.

Raubenheimer J H, 1995. GIS vergemaklik toegang tot Ruimtelike Data. *Proceedings of the Cape Technikon Technology Conference in Civil Engineering*, Cape Town, 9 pages.

Rhind D, 1993. Policy on the supply and availability of Ordnance Survey information over the next five years. *Mapping Awareness*, 7(1 & 2), p.37-41.

Rhind D, 1992. Data access, charging and copyright and their implications for geographical information systems. *International Journal of Geographical Information Systems*, 6(1), p.13-30.

Robertson W A and Aitken G D, 1992. Funding National Mapping. *The Cartographic Journal*, 29(1), p.16-19.

Robson P G and Adlam K A M, 1991. A menu-aided retrieval system (MARS) for use with a relational database management system. In *Metadata in the Geosciences*, Medyckyj-Scott D, Newman I, Ruggles C and Walker D (ed.). Loughborough: Group D Publications.

Rogers B, 1988. The BBC microcomputer-based map library system at Plymouth

Polytechnic. *The Bulletin of the Society of University Cartographers*, 21(2), p.75-81.

Ruggles C and Newman I, 1991. The MRRL's Meta-information Retrieval and Access System. In *Metadata in the Geosciences*, Medyckyj-Scott D, Newman I, Ruggles C and Walker D (ed.). Loughborough: Group D Publications.

Sabel C E and Ralphs M P, 1992. Domesday 2000 - a National Land Information System for Britain. *Mapping Awareness & GIS in Europe*, 6(2), p.55-58.

Salgé F, 1991. Development of Cartographic Databases. In *Advances in Cartography*, Muller J C (ed.). London: Elsevier.

Seaborn D, 1995. Database Management in GIS: is your system a poor relation? *GIS Europe*, 4(5), p.34-38.

Scott D J and Alper S, 1990. Automated chart production system development at NOS. *Proc. ACSM-ASPRS Annual Convention*, Vol. 3, p.249-260.

Shapiro R F and Martin C E, 1989. Challenges and solutions in map publishing. *Proc. ASPRS/ACSM Annual Convention*, Vol. 5, p.237-245.

Smith B, 1994. National Land Information System pilot project. *Mapping Awareness*, 8(10), p.34-38.

Smith D, 1994. Historic maps in the investigation of development sites and the identification of contamination. *The Cartographic Journal*, 31(1), p.3-13.

Smith W E, 1993. Creating a Spatially Accurate Base Map Foundation. *Public Works*, May 1993, p.62-78.

Smits J, 1993. The topographic record of The Netherlands 1800-1992 for sale. *The Cartographic Journal*, 30(1), p.30-39.

Sörbom P and Svensson S, 1990. Plan information in cadastral based Land Information Systems - Possibilities and problems. *Proceedings of the XIX FIG Congress*, p.252-260.

*The STANDARD for data for the National Land Information System*, 1990. Co-ordinating Committee for the National Land Information System.

Starr LE, 1990. Digital Cartography. *Surveying and Land Information Systems*, 50(2), p.93-96.

Steele G F, 1993. The acquisition of maps in digital form in an academic library. *The Cartographic Journal*, 30(1), p.57-61.

Strater S, 1992. Developing a GIS database. *American City and County*, 107, p.16.

Sussman R, 1993. Municipal GIS and the enterprise data model. *International Journal of*

*Geographical Information Systems*, 7(4), p.367-377.

Taylor D R F (ed.), 1991. *Geographic Information Systems: The microcomputer and modern cartography*. Oxford: Pergamon.

Thapa K and Burtch R C, 1991. Primary and Secondary Methods of Data Collection in GIS/LIS. *Surveying and Land Information Systems*, 51(3), p.162-170.

Thorley G, Griffeth M E and Chambers M, 1993. National Mapping Division of the U.S. Geological Survey: building partnerships for map and data production. *Surveying and Land Information Systems*, 53, p.241-243.

Tosta N, 1992. GIS Standards. *Geo Info Systems*, Oct. 1992, p.28-31.

Toth C and Schenk T, 1992. A map-referenced data base system for real-time DEM data acquisition. *Proc. ASPRS-ACSM Annual Convention*, Vol. 2, p.230-236.

Tripp R E and Coats J W, 1988. Hold down the cost of mapping distribution facilities. *Electrical World*, 202, p.54-56.

Van Reij J B, 1990. LIS-GIS, Contrasts and similarities; a commercial approach. *Proceedings of the XIX FIG Congress*, p.330-343.

Visvalingam M, 1991. Cartography and Geographical Information Systems. A Summary

Report by the British Cartographic Society. *Education, Training and Research Publication No. 2, ETH09/91*, Association for Geographic Information (IGO), 10 p.

Volkwyn R, 1995. Why is GIS data so expensive? *Computer Graphics*, 6(3), p.51.

Von Meyer N and Scruggs R, 1990. Cadastral surveyors, their data, and land information systems. *Proc. ACSM-ASPRS Annual Convention*, Vol. 3, p.214-220.

Vrana R, 1989. Historical data as an explicit component of land information systems. *International Journal of Geographical Information Systems*, 3(1), p.33-49.

Walker D R F, Newman I A, Medyckyj-Scott D J and Ruggles C L N, 1992. A system for identifying datasets for GIS users. *International Journal of Geographical Information Systems*, 6(6), p.511-527.

Wan W Y and Williamson I P, 1994. Solutions to maintaining associativity in LIS with particular reference to the needs of the utility industry. *The Australian Surveyor*, 39(4), p.290-296.

Wan W Y and Williamson I P, 1995. The Users' view of Digital Cadastral Databases in Australia. *The Australian Surveyor*, 40(1), p.53-62.

Waters H, 1993. Project management of mapping projects in developing countries in Africa. *The Cartographic Journal*, 30(1), p.52-56.

Welch R, Remillard M and Alberts J, 1992. Integration of GPS, Remote Sensing, and GIS Techniques for Coastal Resource Management. *Photogrammetric Engineering & Remote Sensing*, 58(11), p.1571-1578.

Wilkinson K, 1995. Commercially available data in South Africa. *Computer Graphics*, 6(3), p.56-60.

Woodfin T M and Dignum D, 1991, A pragmatic GIS database design to facilitate economic development in Jefferson County, Texas. *Proceedings URISA '91 conference*, p.166-178.

## **APPENDIX A**

### **QUESTIONNAIRE ON SPATIAL DATA**

University of Cape Town



<p>SECTION A</p> <p>GENERAL</p>
---------------------------------

- 1. Name of person completing the form:  
.....
- 2. Your profession/occupation:  
.....
- 3. Organisation:  
.....
- 4. Tel. (        ) ..... Fax. (        ) .....
- 5. Postal Address: .....  
.....  
..... Code: .....
- 6. Briefly describe the business of your organisation:  
.....  
.....  
.....  
.....

## SECTION B

## MAPS AND SURVEY DATA

Please mark the columns which are applicable to you.

- A Maps/spatial data that you produce and are available to the public.
- B Maps/spatial data that you produce for internal use only.
- C Maps/spatial data used and produced by your organisation.
- D Maps/spatial data used by your organisation but obtained from elsewhere.

## 1. Maps

A	B	C	D
AVAILABLE		OBTAIN	
Public	Int.	Int.	Else.

## 1.1 Topographical / Contour / Detail maps

Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## 1.2 Political and/or administrative maps

Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		A AVAILABLE Public	B Int.	C OBTAIN Int.	D Else.
1.3	Aeronautical Maps				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.4	Geological maps				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.4.2	Mineral rights maps				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.5	Vegetation maps				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.6	Road maps				
1.6.1	Road maps				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		A AVAILABLE Public	B Int.	C OBTAIN Int.	D Else.
1.6.2 Street maps					
Scales:	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.6.3 Tourist maps					
Scales:	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.7 Cadastral maps					
<u>Type</u>	<u>Scale</u>				
Diagrams	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
. . . . .	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
. . . . .	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.8 City maps					
<u>Type</u>	<u>Scale</u>				
Land-use	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zoning	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Engineering	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Building plans	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
. . . . .	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
. . . . .	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		A AVAILABLE Public	B Int.	C OBTAIN Int.	D Else.
1.9	Hydrographic maps				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.10	Historical maps				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.11	Distribution maps				
1.11.1	Statistical maps				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.11.2	Census maps				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.11.3	Population maps				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.12	Weather maps				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		A AVAILABLE Public	B Int.	C OBTAIN Int.	D Else.
1.13.1	Digital maps				
	Vector	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Raster	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.13.2	Type of information shown on the digital maps:				
	.....				
	.....				
	.....				
1.13.3	Digital terrain model (DTM) / DEM	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1.13.4	Does your digital data comply with the National Exchange Standard? .....				
	If not, please give reasons and say if it will be done in future: .....				
	.....				
	.....				
	.....				
	.....				

## 2. Remote Sensing Data

### 2.1 Aerial photos

Type	Scale				
Black/white .....		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Colour .....		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Infrared .....		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
False colour .....		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		A	B	C	D
		AVAILABLE		OBTAIN	
		Public	Int.	Int.	Else.
2.2	Diapositives/negatives				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.3	Orthophotos				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.4	Flight plans				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.5	Remote sensing images				
	Scales: . . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	. . . . .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	Co-ordinates (positions) & heights				
3.1	Co-ordinates of cadastral (property) beacons.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.2	Co-ordinates of trigonometrical beacons or town survey marks.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.3	Co-ordinates of photo control points	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
<b>AVAILABLE</b>		<b>OBTAIN</b>	
<b>Public</b>	<b>Int.</b>	<b>Int.</b>	<b>Else.</b>

3.4 Co-ordinates of points obtained from surveyors or engineers. ☐ ☐ ☐ ☐

Please give details: . . . . .  
 . . . . .  
 . . . . .

3.5 Co-ordinates of construction points ☐ ☐ ☐ ☐

3.6 Are these co-ordinates on the Lo system? Yes ☐ No ☐

3.7 Heights ☐ ☐ ☐ ☐

Please give details: . . . . .  
 . . . . .  
 . . . . .

4. Atlases ☐ ☐ ☐ ☐

Please give details: . . . . .  
 . . . . .  
 . . . . .

5. Other maps or spatial data ☐ ☐ ☐ ☐

Please give details: . . . . .  
 . . . . .  
 . . . . .  
 . . . . .



6. Does your organisation have a catalogue/price list of maps or spatial data? Yes No
- ☐ ☐

If so, please send me a copy.

7. Please give information on the copyright that exist on your data. . . . .
- . . . . .
- . . . . .
- . . . . .
- . . . . .
8. Are there any maps, co-ordinates or spatial data you need but which are not available? . . . . .
- . . . . .
- . . . . .
- . . . . .
- . . . . .
9. Would you supply data for a central database (information system) if you would in turn have access to such database? . . . . .
- . . . . .
- . . . . .
- . . . . .
10. Are you currently using a GIS? . . . . .
11. If not, (a) do you have access to a GIS? . . . . .
- (b) do you intend obtaining a system within the next 2 years? . . . . .
- . . . . .
- . . . . .

**Thank you for your co-operation.**

**APPENDIX B**

**ORGANISATIONS OR INDIVIDUALS  
THAT COMPLETED THE  
QUESTIONNAIRE ON SPATIAL DATA**

Air survey companies

Photogrammetrists

Directorate of Surveys and Land Information:

Director of Mapping

Director of Control Surveys

Map production organisations:

Map Studio and AA Mapping & Publishing

Council for Geoscience (Geological Survey)

Cartographers

Land surveyors

Surveyors

Department of Water Affairs & Forestry

ICL (SA) (Genamap)

Institute for Geographical Analysis  
(University of Stellenbosch)

Provincial Administration of the Western Cape

Wellington Municipality

South African National Defence Force

Civil engineers

Town planners

Town planning technicians

Architects

Landscape architects

Quantity surveyors

Surveyor-General

Deeds Office

Conveyancers

Nature conservation

Ecologists

Computing Centre for Water Research (University of Natal)

Botanists

Elsenburg Agricultural Development Institute

Agricultural Research Institute

Farmers

Property developers

Estate agents

Property valuers

S. A. Library

Telkom

Eskom

Traffic police

Economist (Economic analysis)

Satellite Application Centre (CSIR)

CSIR (Land use Management Services)

Dept. of Surveying & Geodetic Engineering (UCT)

Meteorological Office